

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 4, 2005, 22:43:16 ; Search time 230 Seconds  
(without alignments)  
3251.211 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MWIGNSILLIIFASSIS.....VTPHCTSLSEISDEALYNY 457

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405566

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp  
-O=/cg2\_1/USPTO.spool.p/US10691374/runatc.28022005.120706.2101:/app\_query.fasta\_1.647  
-DB=Issued Patents NA -QFMT=fastop -SUPFIX=ext1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10691374.@CGN1\_1.105.@runatc.28022005.120706.2101 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:\*  
1: /cg2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cg2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cg2\_6/prodata/1/ina/5A\_COMB.seq:\*  
4: /cg2\_6/prodata/1/ina/5B\_COMB.seq:\*  
5: /cg2\_6/prodata/1/ina/FACTUS\_COMB.seq:\*  
6: /cg2\_6/prodata/1/ina/backfillseq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	1636	6	5447867-2
2	2390	100.0	1636	6	5447867-2
3	2390	100.0	1636	6	5447867-2
4	1113	46.6	1631	3	US-09-051-239A-1
5	1103	46.6	1631	4	US-10-151-668-1
6	787	32.9	1657	3	US-08-941-532-5
7	786.5	32.9	1726	3	US-08-467-023-133
8	768	32.1	1479	3	US-08-467-023-140
9	768	32.1	1395	3	US-08-467-023-140
10	366	15.3	1410	3	US-08-467-023-139
11	366	15.3	219	6	5453566-1
12	366	15.3	219	6	5453566-1

13	299	12.5	1356	4	US-09-107-532A-2998	Sequence 2998, App
14	296	12.4	1086	4	US-09-614-221A-573	Sequence 573, App
15	293.5	12.3	1570	4	US-09-787-583-1	Sequence 1, App1
16	241	10.1	1611	1	US-08-061-062A-5	Sequence 5, App1
17	241	10.1	1611	1	US-08-536-150-5	Sequence 5, App1
18	221	9.2	2974	3	US-08-290-978A-4	Sequence 4, App1
19	221	9.2	2974	2	US-08-780-869-4	Sequence 4, App1
20	212.5	8.9	1329	4	US-09-107-532A-2600	Sequence 2600, App
21	205.5	8.6	1348	3	US-09-198-955-5	Sequence 5, App1
22	205.5	8.6	1248	3	US-09-670-141-5	Sequence 5, App1
23	192	8.0	864	3	US-09-087-194-34	Sequence 34, App1
24	189	7.9	164	3	US-08-941-532-7	Sequence 7, App1
25	186	7.8	710	3	US-08-998-416-603	Sequence 603, App
26	169	7.1	3250	1	US-08-061-062A-7	Sequence 7, App1
27	169	7.1	3250	1	US-08-536-150-7	Sequence 7, App1
28	164	6.9	219	3	US-09-051-239A-9	Sequence 9, App1
29	164	6.9	219	3	US-09-151-668-9	Sequence 9, App1
30	146.5	6.1	1664976	4	US-08-916-421B-1	Sequence 1, App1
31	146.5	6.1	1664976	4	US-09-692-570-1	Sequence 1, App1
32	134	5.6	155	3	US-09-051-239A-7	Sequence 7, App1
33	134	5.6	155	3	US-10-151-668-7	Sequence 7, App1
34	129.5	5.4	3018	3	US-09-206-942-40	Sequence 40, App1
35	129.5	5.4	3018	3	US-09-206-942-38	Sequence 38, App1
36	127	5.3	2013	4	US-09-107-433-2373	Sequence 2373, App
37	127	5.3	2127	4	US-09-583-110-358	Sequence 358, App
38	126	5.3	2577	4	US-09-733-643B-1	Sequence 1, App1
39	124.5	5.2	2706	4	US-09-134-001C-2514	Sequence 2514, App
40	123.5	5.2	1779	4	US-09-302-626B-5	Sequence 5, App1
41	121	5.1	1483	3	US-08-961-527-315	Sequence 315, App
42	120.5	5.0	3015	3	US-09-206-942-56	Sequence 56, App
43	120.5	5.0	3033	3	US-09-206-942-54	Sequence 54, App1
44	118.5	5.0	3222	3	US-09-206-942-46	Sequence 46, App1
45	118.5	5.0	3240	3	US-09-206-942-46	Sequence 46, App1

#### ALIGNMENTS

RESULT 1  
5447867-2  
Patent No. 5447867  
APPLICANT: BRIDGES, IAN/SCHUCH, WOLFGANG/GRIERSON, DONALD  
TITLE OF INVENTION: RECOMBINANT DNA CONTAINING PECTIN  
ESTERASE GENE SEGMENTS  
NUMBER OF SEQUENCES: 4  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/24, 866  
FILING DATE: 26-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 720, 629  
FILING DATE: 25-JUN-1991  
APPLICATION NUMBER: 419, 779  
FILING DATE: 29-SEP-1989  
APPLICATION NUMBER: 119, 614  
FILING DATE: 12-NOV-1987  
SEQ ID NO: 2  
LENGTH: 1636  
5447867-2

#### Alignment Scores:

Pred. No.: 3.35e-273  
Score: 2390.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 6  
Length: 1636  
Matches: 457  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-691-374-2 (1-457) x 5447867-2 (1-1636)

QY 1 MetValIleGlnArgAsnSerIleLeuLeuIleIleIlePheAlaSerSerIleSer 20  
DB 51 ATGGTATTCACCAAGAGATGATTCCTTCATCATTAATATTTGGCTTCATCAATTTCA 110  
QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheIleGlnValTyrAspAsnIleLeu 40

Db	111	ACTTGTGAGCGCAAGTATTGATGATGACAAATTATTTCAACAAGTTTATGATTAATATCTT	170
QY	41	GIUINGIUPHEALAHASPPheGlnAlaTYrLeuSerTYrLeuSerLYsAsnIleGlu	60
Db	171	GAACAAGAAATTTGCTCATGATTTTCAAGCTTATCTTTCTTATTGGACCAAAAATATGGAA	230
QY	61	SerAsnAsnAsnIleAspLYsValAspLYsAsnGlyIleYsValIleAsnValLeuSer	80
Db	231	AGCAACAATTAATATTGACAAGGTGGATTAATAATGGATTTAAAGATTTAATGTACTTACG	290
QY	81	PheGlyAlaLYsGlyAspGlyLYsThrTYrZaspAsnIleAlaPheGluGlnAlaTrpAsn	100
Db	291	TTTGAGAGCTTAAGGGTGATGAAAAACATATATATATTTGATTGGACCAAGATGGAAAT	350
QY	101	GIUAlaCYsSerSerArgThrProValGlnPheValAlaProLYsAsnLYsAsnTYrLeu	120
Db	351	GAAGATGTTATCATCTAGAACACCTGTTCAATTTGTGTTCTTAAAAACAAGATTTATCTT	410
QY	121	LeuLYsGlnIleThrPheSerGlyProCYsArgSerSerIleSerValLYsIlePheGly	140
Db	411	CTCAGCAAAATCACCTTTTTCAGGTCATCGAGATCTTCTATTTCAAGTAAAGATTTTGGAA	470
QY	141	SerLeuGlnAlaSerSerLYsIleSerAspTYrLYsAspArgArgLeuTrpIleAlaPhe	160
Db	471	TCTTAAACAACATCTAGTAAATAATTTTCAGACTTAACAAGAATGAAGGCTTTGGATGCTTTT	530
QY	161	AspPheValGlnAsnLeuValAlaGlyGlyGlyGlyThrIleAsnGlyAsnGlyGlnVal	180
Db	531	GATATGCTTCAAAATTTTATGTTTGGAGGGAGGAGAACTTACATGAGCAATGACAAAGTA	590
QY	181	TrpTrpProSerSerCYsLYsIleAsnLYsSerLeuProCYsArgAspAlaProThrala	200
Db	591	TGGTGGCCAAAGTTCTTGCAAAATTAATAATCACTGCATGCAAGGATCCACCAACGGCC	650
QY	201	LeuThrPheTrpAsnCYsLYsAsnLeuLYsValAsnAsnLeuLYsSerLYsAsnIleGln	220
Db	651	TTAACTCTTCGGAATTGCAAAATTTGAAAGTAAATTAATCTTAAAGTAAATATCACAA	710
QY	221	GlnIleHisIleLYsPheGluSerCYsTrpAsnValAlaIleSerAsnLeuMetIleAsn	240
Db	711	CAAAATTCATATCAAAATTTTGAGTCATCACTAATGTTTATGCTTCAAAATTTATGATCAAT	770
QY	241	AlaSerAlaLYsSerProAsnThrAspGlyValHisValSerAsnThrGlnTYrIleGln	260
Db	771	GCTTCAGCAAAAGGCCAAATATCTAGTGGAGTCATGATCAATCAATCAATATATTTCAA	830
QY	261	IleSerAspThrIleIleGlyThrGlyAspAspCYsIleSerIleValSerGlySerGln	280
Db	831	ATATCTGATACTATTATTGGAACAGGTATGATTTATTTCAATGTTCTTCGAGTTCTCAA	890
QY	281	AsnValGlnAlaIleThrasnIleThrCYsGlyProGlyHisGlyIleSerIleGlySerLeu	300
Db	891	AATGTGACGGCCACAATTAATTACTGTGTGTCACAGTCACTGATTAATTAATTTGAAGCTTA	950
QY	301	GlySerGlyAsnSerGlyAlaTYrValSerAsnValTrpValAsnGlnAlaLYsIleIle	320
Db	951	GGATCTGGAAATTCAGAGCTTATGTGTCTATGTCTTACTGTAAATGAAGCCAAAATTAATC	1010
QY	321	GlyAlaGluAsnGlyValArgIleLYsThrTrpGlnGlyGlySerGlyGlnAlaSerAsn	340
Db	1011	GGTGCCGAAATGGAGTTAGATCAAGACTTCGACGGAGAGATCTCGCAACGTAAGCAAC	1070
QY	341	IleLYsPheLeuAsnValGluMetGlnAspValIleLYsTYrProIleIleIleAspGlnAsn	360
Db	1071	ATCAAAATTTTCGAGATGTGAATGCAAGACGTAAAGTATCCCAATTAATTAATTAAGCCAAAC	1130
QY	361	TYrCYsAspArgValGluProCYsIleGlnGlnPheSerAlaValGlnValLYsAsnVal	380
Db	1131	TATGTGATCAGTGTGAACCATGTATCAACAGTTTCAAGAGTTCAAGTGAAGAAATGTG	1190
QY	381	ValTYrGluAsnIleLYsGlyThrSerAlaThrLYsValAlaIleLYsPheAspCYsSer	400

Db	1191	GTGATATGAAATATCAAGGGCCAAAGTCGACAAAGAGTGGCCATATAATTTGATTGGACG	1250
Qy	401	ThrasnheProcysgluyllelleMetGluenilleaenleuValIglYlserGly	420
Db	1251	ACAACTTTCATGTGAGGAATTTAATATGAGAAATTAATTTAATTTAGTGGGAAAGTGG	1310
Qy	421	lyeProSerGluAlaThrCyselysAsnValHisPheAsnAsnAlaGluHisValThrPro	440
Db	1311	AAACCATCAGAGGCTACGGTCAAAAATGCTCAATTTTAAACAATGCTGAACAATGTTACCA	1370
Qy	441	HisCysThrSerleuGluIleSerGluAspGluAlaLeuIleuTyraAsnTy	457
Db	1371	CACGTGACCTTCACCTAGAAATTTCAAGAGATGAGCTTTTGTATTAATTAAT	1421
RESULT 2			
5447867-2			
Patent No. 5447867			
APPLICANT: BRIDGES, IAN; SCHUCH, WOLFGANG; GRIERSON, DONALD			
TITLE OF INVENTION: RECOMBINANT DNA CONTAINING PECTIN			
ESTERASE GENE SEGMENTS			
NUMBER OF SEQUENCES: 4			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/24,866			
FILING DATE: 26-FEB-1993			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 720,629			
FILING DATE: 25-JUN-1991			
APPLICATION NUMBER: 419,779			
FILING DATE: 29-SEP-1989			
APPLICATION NUMBER: 119,614			
FILING DATE: 12-NOV-1987			
SEQ ID NO:2:			
5447867-2			
ALIGNMENT SCORES:			
Pred. No.: 3,35e-273			
Score: 2390.00			
Percent Similarity: 100.00%			
Best Local Similarity: 100.00%			
Query Match: 100.00%			
DB: 6			
Gaps: 0			
US-10-691-374-2 (1-457) x 5447867-2 (1-1636)			
Qy	1	MetValIleGlnAyrAsnSerIleLeuLeuIleIleIlePheAlaSerSerIleSer	20
Db	51	ATGGTATCCAAAGAAATAGTATTCCTTCATTAATTAATTTGGCTTCATCAATTTCA	110
Qy	21	ThrCysArgSerAsnValIleAspAsnAsnLeuPheGluValTyraAspAsnIleLeu	40
Db	111	ACTGTGAGAACATGTATTATGTATGACAAATTTATTCAAACAAGTTATGATTAATTTCTT	170
Qy	41	GluGlnIleuPheAlaHisAspPheGlnAlaTyLeuSerTyLeuSerTyAsnIleGlu	60
Db	171	GAAACAATAATTTCTCATGATTTTCAAGCTTATCTTTCTTATTTGAGCAAAAAAATTTGAA	230
Qy	61	SerAsnAsnAsnIleAspIysValAspIysAsnGlyIleIysValIleAsnValLeuSer	80
Db	231	AGCAACAATATATTTGACAAAGTTGATATAAAATGGATTTAAAGTATTAATGTACTTAC	290
Qy	81	PheGlyValAlaIysGlyAspGlyIysThrTyraAspAsnIleAlaPheGluGlnAlaTrpAsn	100
Db	291	TTTGGAGCTTAAGGGTATGAAAAAACATATATATATTTGCAATTTGAGCAAGCATGGAAT	350
Qy	101	GluAlaCysSerSerArgThrProValGlnPheValAlaPolyAsnIysAsnTyLeu	120
Db	351	GAAAGATGTTTCACTTCAAGAACACCTGTTCAATTTGTGTCTTAAACAAAGAAATATCTT	410
Qy	121	IleuIysGlnIleThrPheSerGlyProCysArgSerSerIleSerValIysIlePheGly	140
Db	411	CTCAAGCAATACCTTTTCAGGTTCACGATCTTCATTTTCAGTAAAGATTTTGG	470

```
QY 141 SerLeuGluAlaSerSerLysIleSerAspTyrLysAspArgArgLeuTrpIleAlaPhe 160
DB 471 TCTTATGAAGCATCTTGATTAATTTGAGATACAAAGATGAAGGCTTTGATTCCTTT 530
QY 161 AspSerValGlnAsnLeuValValGlyGlyGlyThrIleAsnGlyAsnGlyGlnVal 180
DB 531 GATAGAGTTCAAAATTTAGTTGTGGAGAGGAGGAACTATCATATGGCATGGACAAAGTA 590
QY 181 TrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaProTrpAla 200
DB 591 TGGTGCCCAAGTTCTTGCAAAATTAATCACTCCATCGACGGATGCACCAACGGCC 650
QY 201 LeuThrPheThrAsnCysLysAsnLeuValAsnAsnLeuLysSerLysAsnAlaGln 220
DB 651 TTAACCTTCTGAAATGGCAAAATTTGAAAGTAATTAATCTAAAGATGAAGTAATGACAA 710
QY 221 GlnIleHisIleLysPheGluSerCysThrAsnValAlaIleSerAsnLeuMetIleAsn 240
DB 711 CAATTCATATCAAAATTTGATGTCATGCACTAATGTTGAGCTTCAAAATTTGATGATCAAT 770
QY 241 AlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260
DB 771 GCTTCGCAAAAGCCCAATCACTGATGAGTCATCATGATCAAAATCTCAATATATTCAA 830
QY 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280
DB 831 ATATCTGATCTATTTATGGACAGGTGATGATGATTTCAATTTCTTGATCTCA 890
QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300
DB 891 AATGTCGAGCCCAAAATATTCTTGTCGACGATCAAGTATTAAGTTTGGAACTTA 950
QY 301 GlySerGlyAsnSerGlyAlaTyrValSerAsnValThrValAsnGluAlaLysIleIle 320
DB 951 GGAATCTGGAATTCACAAAGCTTATGCTTAAGTAACTGTAATTAAGCAAAATATTATC 1010
QY 321 GlyAlaGluAsnGlyValArgIleLysThrTrpGlnGlyGlySerGlyGlnIleSerAsn 340
DB 1011 GGTGCCAAATGGAATGAGATCAAGATCTGGCAGAGAGATCTGACAAAGCTTACCAAC 1070
QY 341 IleLysPheLeuAsnValGlnMetGlnAspValLysTyrProIleIleIleAspGlnAsn 360
DB 1071 ATCAAAATTTCTGAATGTGGAATGCAAGACGTTAAGTATCCCATTAATTATAGCCAAAC 1130
QY 361 TyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValLysAsnVal 380
DB 1131 TATGTGATCGAGTTGAAACCATGATATCAACAGTTTTCAGCAGTTCAAGGAAAAATGTG 1190
QY 381 ValTyrGluAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAspCysSer 400
DB 1191 GTGATGAAATATCAAGGCAACAGTCACAAAGGTGGCCATTAATTTGATTCACG 1250
QY 401 ThrAsnPheProCysGlyGlyIleIleMetGluAsnIleAsnLeuValGlyGluSerGly 420
DB 1251 ACAAACTTCCATGTGAAGAAATTAATAGAGAAATTAATTAATTAATTAATTAATTAATTA 1310
QY 421 LysProSerGluAlaThrCysLysAsnValHisPheAsnAsnAlaGlnHisValThrPro 440
DB 1311 AAACCACTCAGAGGCTTCGTCGCAAAATGTCATTTTACAAATGCTTCAACATGTTTACACA 1370
QY 441 HisCysThrSerLeuGlnLysSerGluAspGluAlaLeuLeuTyrAsnTyr 457
DB 1371 CACTGCACCTTCACTAGAAATTTTCAGAGATGAAGCTCTTTGATTAATTAAT 1421
```

```
APPLICANT: BORKHARDT, Bernard
APPLICANT: SANDER, Lili
APPLICANT: PETERSEN, Morten
APPLICANT: BUNDGARD, POULSEN, Gert
APPLICANT: BOTTERMAN, Johan
TITLE OF INVENTION: Seed Shattering
FILE REFERENCE: 2121-0138P
CURRENT APPLICATION NUMBER: US/09/051,239A
CURRENT FILING DATE: 1998-09-28
PRIOR APPLICATION NUMBER: PCT/EP96/04313
PRIOR FILING DATE: 1996-10-04
PRIOR APPLICATION NUMBER: EP 95 402241.4
PRIOR FILING DATE: 1995-10-06
PRIOR APPLICATION NUMBER: EP 95 203328.0
PRIOR FILING DATE: 1995-12-08
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 1
LENGTH: 1631
TYPE: DNA
ORGANISM: Brassica napus
FEATURE:
OTHER INFORMATION: Location 95-163 = region encoding the presumed
OTHER INFORMATION: endo-PG signal peptide.
OTHER INFORMATION: Location 884-900 = region of the endo-PG cDNA
OTHER INFORMATION: corresponding to oligonucleotide PG3
OTHER INFORMATION: Location 1059-1073 = region of the endo-PG cDNA
OTHER INFORMATION: complementary to oligonucleotide PG2
OTHER INFORMATION: Location 1229-1245 = region of the endo-PG cDNA
NAME/KEY: CDS
LOCATION: (95)..(1393)
OTHER INFORMATION: Location 821-837 = region of endo-PG cDNA
OTHER INFORMATION: corresponding to oligonucleotide PGI.
OTHER INFORMATION: Strain cv. Topaz.
NAME/KEY: unsure
LOCATION: (1439)
OTHER INFORMATION: n = a, c, g, t, any, other, unknown, or other
US-09-051-239A-1
Alignment Scores:
Pred. No.: 1,3e-121 Length: 1631
Score: 1113.00 Matches: 223
Percent Similarity: 65.32% Conservative: 69
Best Local Similarity: 49.89% Mismatches: 121
Query Match: 46.57% Indels: 34
DB: 3 Gaps: 8
US-10-691-374-2 (1-457) x US-09-051-239A-1 (1-1631)
QY 7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
DB 119 GCTGTTTCTTATGCGTCTTTTGATGTCGCTTGTGCAAGCTTTGAGTGAAGCAACGTA 178
QY 27 IleAspAsnLeuPheLysGlnValTyrAspAsnIleLeuGluGlnGluPheAlaHis 46
DB 179 ---GATGATGCA-----TATGCTCAT 196
QY 47 AspPheGlnAlaTyrLeuSerTyrLeuSerLysAsnIleGluSerAsnAsnAsnIleAsp 66
DB 197 GAAAGATGCA-----AGCTTCGAAATCCGATGATTAAATCAAGTCACAAACGACGAC 247
QY 67 LysVal-----AbpLysAsnGlyIleLysValIleAsnVal 78
DB 248 GACGTTCTTACCTTGAAAGCTCTGATAGACCCACTACCGAATCATCACTGTACTGT 307
QY 79 LeuSerPheGlyAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGluGlnAla 98
DB 308 TCGAACTTCGAGCCCAAGAGATGGAAGAAACCGATGATCTCAGGCTTTCAAGAAAGCA 367
QY 99 TrpAsnGluAlaCysSerSerArgThrProValGlnPheValValProLysAsnLysAsn 118
DB 368 TCGAAGAGGATGTTCAACAAATGAGATTAAGTACTTCTTAAATCTTAATTCCTTAAGAAAGACT 427
```

```

Qy 119 TyrlleuLeuLysGlnIleThrPheSerGlyProCysArgSerIleSerValLysIle 138
Db 428 TATCTCTTAAGCTTATTAGATTACAGAGGCCATGCAATCTTACTAGCTTCCAGATC 487
Qy 139 PheGlySerLeuGlnLysSerIleSerIleSerIleSerIleSerIleSerIle 157
Db 488 CTAGGCACTTACAGCTTCTTACAAACGATCGATTCAGTAATGACAAAGCAACCTG 547
Qy 158 ILeAlaPheAspSerValGlnLeuValLysGlyGly-----GlyThrIleAsn 175
Db 548 CTTAATTTGGAAACCTTAATATCTAATCAATCGATGGCGCTCGCGCGGATTTGAT 607
Qy 176 GlyAsnGlyGlnValIleThrProSerSerGlyIleAsnLysSerIleProCysArg 195
Db 608 GGAACGGAATATCTCGTGGCAAACTCATGCAAAATCGAACAATCTTAACCATGACA 667
Qy 196 AspAlaProThrAlaLeuThrPheTrpAsnGlyLysAsnLeuValAsnAsnLeuLys 215
Db 668 AAAGCCCAACGGCTTCTACTCTTACCAACCTTAAGATTTGATGTAAGAAATCTGGA 727
Qy 216 SerLysAsnAlaGlnGlnIleHisIleLysPheGlySerGlyThrAsnValAlaLaser 235
Db 728 GTGAGAAATCGACAGAGATTCAGATTGCAATTGGAATGCAACAATGTTGGCGTTAAG 787
Qy 236 AsnLeuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsn 255
Db 788 AATGTTAAGATCATCTGCTCTGGCGATGTCCTCCAAACAGATGTTATTCATTCGTTGCT 847
Qy 256 ThrGlnIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIle 275
Db 848 ACTAAAAACATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 907
Qy 276 ValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyValGlyIle 295
Db 908 GAGATGATCGCAAAATGTTCAATCAATGATTTACTTGGCGCCCGCTCAATGAGATC 967
Qy 296 SerIleGlySerLeuGlySerGlyAsnSerGlyAsnValIleValSerAsnValIleAsn 315
Db 968 AGCATTTGAGAGCTTGGGGGATGACAAATTCGAAAGCTTATGTATCGGGAATGATGAT 1027
Qy 316 GlnAlaLysIleIleGlyAlaGlnAsnGlyValArgIleLysThrTrpGlnGlyGlySer 335
Db 1028 GGTCTCAGCTCTCTGAGCTGACATGAGTAAGATCAAGCTTAACTCAAGAGGAGTCA 1087
Qy 336 GlyGlnAlaSerAsnIleLysPheLeuAsnValGlnMetGlnAspValLysTrpProIle 355
Db 1088 GGAACCTGTAAGAACTTAATTCGAAACATTCGTAATGATGATGATGATGATGATG 1147
Qy 356 IleIleAspGlnAsnTrpCysAspArgValGlnProGlyIleGlnGlnPheSerAlaVal 375
Db 1148 ATATCGACCAAGACATCTGCGCAAG--GACAAATGCGAACACCAAGATCTCGGCTT 1204
Qy 376 GlnValLysAsnValValIleTrpGlnAsnIleLysGlyThrSerAlaThrLysValAlaIle 395
Db 1205 CAAGTGAACAAATGCTGTATCAGAACTAAAGGTAAGAGCGCAACGATGCGCATTA 1264
Qy 396 LysPheAspCysSerThrAsnProCysGlnGlyIleIleMetGlnAsnIleAsnLeu 415
Db 1265 ATGTTAATTCAGCTGTAATATCTCAAGGTAATGCTGTAATGTAATGTAATGTAATC 1324
Qy 416 ValGlyLysSerGlyLysProSerGlnAlaThrCysLysAsnValHisPheAsnAla 435
Db 1325 AAAGGA-----GGAAGCTTCTTGGAAATGTCATGTTAAGGATAA 1369
Qy 436 GlnHisValIleThrProHisCys 442
Db 1370 GGCACCTTTCTCTCAAAATGC 1390

```

RESULT 4  
 US-10-151-668-1  
 ; Sequence 1, Application US/10151668  
 ; Patent No. 6797861

```

; GENERAL INFORMATION:
; APPLICANT: ULVSKOV, Peter
; APPLICANT: CHILD, Robin
; APPLICANT: VAN ONCKELIN, Henri
; APPLICANT: PRINSEN, Els
; APPLICANT: BORKHARDT, Bernard
; APPLICANT: SANDER, Lilili
; APPLICANT: PETERSEN, Morten
; APPLICANT: BONDARD POTJSEN, Gert
; APPLICANT: BOTTERMAN, Johan
; TITLE OF INVENTION: Seed Shattering
; FILE REFERENCE: 2121-0138P
; CURRENT APPLICATION NUMBER: US/10/151,668
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US/09/051,239
; PRIOR FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: PCT/EP96/04313
; PRIOR FILING DATE: 1996-10-04
; PRIOR APPLICATION NUMBER: EP 95 402241.4
; PRIOR FILING DATE: 1995-10-06
; PRIOR APPLICATION NUMBER: EP 95 203328.0
; PRIOR FILING DATE: 1995-12-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1631
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: Location 95-163 = region encoding the presumed
; OTHER INFORMATION: endo-PG signal peptide.
; OTHER INFORMATION: Location 884-900 = region of the endo-PG cDNA
; OTHER INFORMATION: corresponding to oligonucleotide PG3
; OTHER INFORMATION: Location 1059-1073 = region of the endo-PG cDNA
; OTHER INFORMATION: complementary to oligonucleotide PG2
; NAME/KEY: CDS
; LOCATION: (95)..(1393)
; OTHER INFORMATION: Location 1229-1245 = region of the endo-PG cDNA
; OTHER INFORMATION: complementary to oligonucleotide PG5
; OTHER INFORMATION: Location 821-837 = region of endo-PG cDNA
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1439)
; OTHER INFORMATION: Strain cv. Topaz.
; OTHER INFORMATION: n = a, c, g, t, any, other, unknown, or other
US-10-151-668-1

Alignment Scores:
Pred. No.: 1,3e-121 Length: 1631
Score: 1113.00 Matches: 223
Percent Similarity: 65.32% Conservative: 69
Best Local Similarity: 49.89% Mismatches: 121
Query Match: 46.57% Indels: 34
Gaps: 8

US-10-691-374-2 (1-457) x US-10-151-668-1 (1-1631)
Qy 7 SerIleLeuLeuLysIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
Db 119 GCTGTTTCTTAAGCTTATTAGATTACAGAGGCCATGCAATCTTACTAGCTTCCAGATC 178
Qy 27 IleAspArgAsnLeuPheLysGlnValIleThrAsnIleLeuGlnGlnIlePheAlaHis 46
Db 179 ---GATGATGCA-----TATGTCAT 196
Qy 47 AspPheGlnAlaIleLysSerTrpLeuSerLysAsnIleGlnSerAsnAsnIleAsp 66
Db 197 GAAGATGGA-----AGCTTGAATCCGATTAATTCAGCTCAAGCTCAACAGACGAC 247
Qy 67 LysVal-----AspLysAsnGlyIleLysValIleAsnVal 78

```

Db 248 GAAGTTCTTACCTTGAAAGCTCTGATAGACCCACCTACCGAATCATCACTGTACTGTT 307  
Qy 79 LeuSerPheGlyAlaIysGlyAspGlyIleThrTyrAspAsnIleAlaPheGluGlnAla 98  
Db 308 TCGAAGCTTCGAGCCAAAGAGATGAGAAAAACCGATGATATCTACGAGCTTCAAGAAAGCA 367  
Qy 99 TrpAsnGlnAlaCysSerSerArgThrProValGlnPheValValProIysAsnIysAsn 118  
Db 368 TGGAGAGACGACATGTTCAACAAATGAGATTACTTCTTAATTCCTTAAGGAAAGACT 427  
Qy 119 TyrLeuLeuGlnIleThrPheSerGlyProCysArgSerIleSerValIysIle 138  
Db 428 TATCTCCTTAAGTATAGATTAGATTCAGAGCCCATGCAAACTTACTAGTGTCTCCAGATC 487  
Qy 139 PheGlySerLeuGlnAlaSerSerIysIleSerAspTyr---LysAspArgArgLeuTrp 157  
Db 488 CTAGGCACTTATACACTTCTACAAACATCGGATTACAGTAAATGACAGAACCACTGAG 547  
Qy 158 IleAlaPheAspSerValGlnAsnLeuValValGlyGly-----GlyThrIleAsn 175  
Db 548 CTTATTTTGGAGACGCTTAATAATTTATCAATCGATGCGGCGCGGATGTTGAT 607  
Qy 176 GlyAsnGlyGlnValTrpTrpProSerSerCysValIleAsnIysSerLeuProCysArg 195  
Db 608 GGCACGCGAAATATCTGCTGCGCAAACTCATGCAAAATGACAAATCTAAGCCATCACA 667  
Qy 196 AspAlaProThrAlaLeuThrPheTrpAsnCysValAsnLeuValAsnLeuIys 215  
Db 668 AAGCGCCACGCGCTCTACTCTCTACAACTTAAGAAATTTGAAATGTAAGAAATCGAGA 727  
Qy 216 SerIysAsnAlaGlnGlnIleIleIleIysPheGlySerCysThrAsnValAlaAlaSer 235  
Db 728 GTGAGAAATGCAAGCAGATTCAGATTCGATTAAGAAATGCAACATTTGCGCTTAAG 787  
Qy 236 AsnLeuMetIleAsnAlaSerAlaIysSerProAsnThrAspGlyValIleValSerAsn 255  
Db 788 AATGTTAAGATCATCTGCTCGCGCATGTCGCCAACGAGTGTGATCATATCGTTGCT 847  
Qy 256 ThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIle 275  
Db 848 ACTAAACCAATCGAATCTCCAAATTCAGACATTGGAGCAGTGATATGTATATCATAT 907  
Qy 276 ValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyAsnGlyIle 295  
Db 908 GAGGATGATCGAAAAATGTTCAATCAATGATTAATTCGCGCCCGCTCATGGGATC 967  
Qy 296 SerIleGlySerLeuGlySerGlyAsnSerGlyAlaIleTyrValSerAsnValThrValAsn 315  
Db 968 AGCATTGGAAGCTTGGGGATGATCAATTCCAAGCTTATGTATCGGGAAATGATGTGAT 1027  
Qy 316 GluAlaIysIleIleGlyAlaGluAsnGlyValArgIleIysThrTrpGlnGlyIysSer 335  
Db 1028 GGTGCTACCTCTCTGAGCATGACATGAGTAAAGATCAAGACTTACCGAGGAGGATCA 1087  
Qy 336 GlyGlnAlaSerAsnIleIysPheLeuAsnValGluMetGlnAspValIysTyrProIle 355  
Db 1088 GGAAGCTCTTAAGAACCTTAATTCAAAACATTCGATGATATATCAAGAAATCCGATC 1147  
Qy 356 IleIleAspGlnAsnTyrCysAspArgValGluProCysIleGlnGlnPheSerAlaVal 375  
Db 1148 ATATATGACACGAACACTGCGCACAAG---GACAATTCGACACGACAGAAATCTGGGCTT 1204  
Qy 376 GlnValIysAsnValValIysGluAsnIleIysGlyThrSerAlaThrIysValAlaIle 395  
Db 1205 CAAGTGACAAATGTCGTATATCAGAACATTAAGATGACGACCAACAATGTGGCGAT 1264  
Qy 396 LysPheAspCysSerThrAsnPheProCysGluGlyIleIleMetGlnAsnIleAsnLeu 415  
Db 1265 ATGTTAATTTGCGGTGAAATTCATTCGAAGATTTGCTTGAAGATGGAACATC 1324  
Qy 416 ValGlyGlySerGlyLysProSerGlnAlaThrCysValAsnValIleAspAsnAsnAla 435  
Db 1325 AAGAGA-----GGAAAGCTTCTTGGAAATGTCAATGTTAAGATTA 1369

Qy 436 GluIleValThrProHisCys 442  
Db 1370 GGCAGCTTTCTCTTAATATGC 1390

RESULT 5  
US-08-941-532-5  
Sequence 5, Application US/08941532  
Patent No. 6096946  
GENERAL INFORMATION:  
APPLICANT: ROBERTS, Jeremy Alan  
APPLICANT: COUP, Simon Allan  
APPLICANT: JENKINS, Elizabeth Sarah  
TITLE OF INVENTION: CONTROL OF POD DENSIFICATION  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/941,532  
FILING DATE: 30-SEP-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB96/00757  
FILING DATE: 29-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9506684.1  
FILING DATE: 31-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Remond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0623.0580001/RWE  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1657 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 145..1446  
US-08-941-532-5

Alignment Scores:  
Pred. No.: 2.05e-120  
Score: 1103.00  
Percent Similarity: 65.54%  
Best Local Similarity: 49.77%  
Query Match: 46.15%  
DB: 3  
Gaps: 7

US-10-691-374-2 (1-457) x US-08-941-532-5 (1-1657)

Qy 7 SerIleLeuLeuLeuIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26  
Db 169 GCTATTTCTTATGCGTTCTTGTATGCTCGCTGCGCAAGCTTTGATGATGCAACGTA 228  
Qy 27 IleAsp-----AspAsnLeuPheIysGlnValTyrAspAsnIleGlu 41  
Db 229 GATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282

```

QY 42 GlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerLysAsnIleGluSer 61
Db 283 CTCAACACGACGACGACGCTTCTTACCTGAAAAAGCTCCGATGACCCACTACCGCATCA 342
QY 62 AsnAsnAsnIleAspLysValAspLysAsnGlyTyrLeuValIleAsnValLeuSerPhe 81
Db 343 TCA-----ACTGTTGCTGTTTCGAACCTTC 366
QY 82 GlyAlaLysGlyAspGlyLysTyrTyrAspAsnIleAlaPheGluGlnAlaTyrAsnGlu 101
Db 367 GAGGCAAAAGGTGATGAGAAAAACCGATGATACAGGCTTTCAGAAACACGGAAGAG 426
QY 102 AlaCysSerSerArgThrProValGlnPheValProLysAsnLysAsnTyrLeuLeu 121
Db 427 GCATGTTTCACAAATGAGAGTACTTCTTGTATCTTAAAGGAGAGCTTATCTTCCTT 486
QY 122 LysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGlySer 141
Db 487 AAGCTATTAAGATTCAAGAGCCCATCCAAATCATTAAGTACGTTCAATCTTCAGGCTTATTTG 546
QY 142 LeuGluAlaSerSerLysIleSerAspTyr---LysAspArgLysLeuTyrIleAlaPhe 160
Db 547 TTATCGCTTTCACAAACGATGAGATTAAGTAAATGACAAACCACTGCTTATTTG 606
QY 161 AspSerValGlnAsnLeuValAlaGlyGly-----GlyThrIleAsnGlyAsnGly 178
Db 607 GAGGAGCTTAATATCATATCATGATGCGGCTCGCGGGGAGTTTGTATGAGCAGCA 666
QY 179 GlnValTyrTyrProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaPro 198
Db 667 AAAATCTGTCGCAAAACCTCATGCAAAATCGACAAATCTAAGCCATCCACAAAAGCCCA 726
QY 199 ThrAlaLeuThrPheTyrAsnCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsn 218
Db 727 ACGGCTTCTTCTCTACACCTTAAACATTTGATGATGAAAGATCTGAGAGTGAAGAAAT 786
QY 219 AlaGlnGlnIleHisIleLysPheGluSerCysThrAsnValAlaAlaSerAsnLeuMet 238
Db 787 GCACGACGATTCAGATTCATTCATGAGAAATGCAACAGTGTATGATTAAGATGTTAAG 846
QY 239 IleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyr 258
Db 847 ATCACTGCTCTCGCGATGATGCCAACACGAGTGGTATTAATCTCTCTCTCTAAAAAC 906
QY 259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly 278
Db 907 ATTGCAATCTCAATTCAGACATTCGAGACAGTGAATGATGATCCATTCAGAGATGA 966
QY 279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyLysSerIleGly 298
Db 967 TCCGAAATGTTCAATCATATCATATCTTAACTTGGGCCCCCGGTCATGCGCATCGCA 1026
QY 299 SerLeuGlySerCysLysSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaLys 318
Db 1027 AGCTTGGGGGATGACAAATTCAAAGCTTATGATGCGGAATTAATGATGATGCTCTACG 1086
QY 319 IleIleGlyAlaGlnAsnGlyValArgIleLysTyrTrpGlnGlyLysSerGlnAla 338
Db 1087 CTCTCGAGACGACAGATGAGATTAAGATCAACATCTTACACAGGAGGCTCGAGAACTGCT 1146
QY 339 SerAsnIleLysPheLeuAsnValAlaLysGlnAspValLysTyrProIleIleIleAsp 358
Db 1147 AAGAAATTAATTAATTCAAACAAATTCGTATGATATTCAGAAATCCGATCATTAATGAC 1206
QY 359 GlnAsnTyrCysAspArgValGlnProCysIleGlnGlnPheSerAlaValGlnValLys 378
Db 1207 CAGAACTACTGCGAAG---GACAAATCCAAACAAAGAAATCTGCGTTCAGGTGAAC 1263
QY 379 AsnValValTyrGlnAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAsp 398
Db 1264 AATGTCGTATCGGAACATCAAGGTACGAGCCAAAGGATGTCGGATTAATCTTTAAT 1323

```

```

QY 399 CysSerThrAsnProCysGlnGlyIleIleMetGluAsnIleAsnLeuValGlyGlu 418
Db 1324 TGCAGTGTGAATATCATCTCCAAAGTATTTGCTTGAGAAATGTGAACATCAAGGA--- 1380
QY 419 SerCysLysProSerGlnAlaThrCysLysAsnValHisPheAsnAsnAlaGlnHisVal 438
Db 1381 -----GAAAAGCTTCTTGCAAAATGTCAATGTTAAGATTAAGCAACCGTT 1428
QY 439 ThrProHisCys 442
Db 1429 TCTCTTAATATGC 1440

RESULT 6
US-08-467-023-133
; Sequence 133, Application US/08467023
; Patent No. 6090386
; GENERAL INFORMATION:
; APPLICANT: Griffith, Irwin J.
; APPLICANT: Bond, Joanne;
; APPLICANT: Garman, Richard D;
; APPLICANT: Kuo, Mei-Chang;
; APPLICANT: Jeung, Su-mei H.;
; APPLICANT: Brauer, Andrew;
; APPLICANT: Exley, Mark A.;
; APPLICANT: Powers, Steven P.
; TITLE OF INVENTION: Allergenic Proteins And Peptides From
; TITLE OF INVENTION: Japanese Cedar Pollen
; NUMBER OF SEQUENCES: 261
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Immunologic Pharmaceutical Corporation, Inc.
; STREET: 610 Lincoln St
; CITY: Waltham
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,023
; FILING DATE: June 6, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,225
; FILING DATE: December 6, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 133:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..1586
; US-08-467-023-133

Alignment Scores:
Pred. No.: 7.15e-83 Length: 1726
Score: 787.00 Matches: 168
Percent Similarity: 58.07% Conservative: 73
Best Local Similarity: 40.48% Mismatches: 160
Query Match: 32.93% Indels: 14

```

```

DB: 3 Gaps: 7
US-10-691-374-2 (1-457) x US-08-467-023-133 (1-1726)
QY 45 AAlHAspPheGlnAlaTyrLeu-----SerTyrLeuSerIysAsnIleGlnSer 61
DB 105 GCGAAGATCAATCTGCCAAATATTATGTTGACAGCTTGTGCAAAAATATCTTAATG 164
QY 62 AenAsnAsnIleAspIysValAsp-----LysAsnGlyIleIysValIleAsnVal 78
DB 165 AATCGAGATTAAAGAAAGTTGAGCATCTTCGTCAATGCTTAATCAACCTTCATG 224
QY 79 LeuSerPheGlyValIysGlyAspGlyIysThrTyrAspAsnIleAlaPheGlnGlnAla 98
DB 225 GAAAGATATGGCCAGTGGCAGTGAAGACATGATTGCACTGAGGCAATTTTCAACAGCA 284
QY 99 TrpAsnGlnAlaCysSerSerArgThrProValGlnPheValAlaProIysAsnIysAsn 118
DB 285 TGGCAAGCTGCATGCAAAAACCATCA---GCAATGTGCTTGTGCGACAGCAAGAA 341
QY 119 TyrLeuLeuIysGlnIleThrPheSerGlyProCysArgSerSerIleSerValIysIle 138
DB 342 TTTGTGTAAACATCTGTTCTTCAATGGCCATGTCATCACTTTCCTTTAAGGTA 401
QY 139 PheGlySerLeuGlnAlaSerSerIysIleSerAspTyrIysAspArgTrpIle 158
DB 402 GATGGATTAATACCTGCTACCAAAATCCAGCGAGCTGGAGAAATTAATGATATG 461
QY 159 AlaPheAspSerValGlnAsnLeuValIleGlyIleGlyIleThrIleAsnGlyAsnGly 178
DB 462 CAGTTTGCTAACTTAACAGTTTACTCTAATGGGTAAAGGTAAATGATGAGCAAGGA 521
QY 179 GlnValIleTyrProSerSerCysIys---IleAsnIysSerLeuProCys-----Arg 195
DB 522 AAACAATGGGGCTGGCCCAATGTAAATGGGTCAATGAGACAGAAATTTGCCAAGATCG 581
QY 196 AspAlaProThrAlaLeuThrPheTrpAsnCysIysAsnLeuIysValIleAsnLeuIys 215
DB 582 GATAGACCAACAGCTTAATTAATTCGATTTTCCAGCGGCTGTAAATTCAGAGCATGAA 641
QY 216 SerIysAsnAlaGlnIleIleHisIleIysPheGlnSerCysThrAsnValAlaIleSer 235
DB 642 CTAAATGACAGTCCGCAATTTCTTAAGTTTGGGAATTTGAGGAGGATGAAATATC 701
QY 236 AenLeuMetIleAsnAlaSerAlaIysSerProAsnThrAspGlyValHisIleSerAsn 255
DB 702 GGCATTAGTATTCGGCAGCAGAGACAGATCTTAACATGATGATGATATCTTGCA 761
QY 256 ThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIle 275
DB 762 TCTAAAACTTTCACTTAACAAAAGAACAGATGAGAACAGGGATACGCGCTATTA 821
QY 276 ValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIle 295
DB 822 GGGACAGGGGCTCTTAATATTTGATGATGAGATCTGATTTGGCGTCAAGCCATGGA 881
QY 296 SerIleGlySerLeuGlySerGlyAsnSerGlnAlaIysValSerAsnValThrValAsn 315
DB 882 AGTATGAGAACTTCTGGAGGAGAAACTCTAGAGCAGAGTTTCATACGTCGACCTAAAT 941
QY 316 GlnAlaIysIleIleGlyAlaGlnAsnGlyValArgIleIysThrTrpGlnGlyIysSer 335
DB 942 GGGGCTTAATTTCTTAACACAAATGAGATTAAAGATCAAAACATGCGGGGTGTCA 1001
QY 336 GlnGlnAlaSerAsnIleIysPheLeuAsnValGlnMetGlnAspValIysTyrProIle 355
DB 1002 GGCATGGCAAGCATATTAATTTATGAGATGTTGAATATGATTAATTCGAGAAACCCAT 1061
QY 356 IleIleAspGlnAsnTyrCysAspArgValGlnProCysIleGlnGlnPheSerAlaVal 375
DB 1062 TTAATTAATCAATTTCTTAACATGCACTTCTGCTTGGCCAAACCAAGAGTCTGGG 1121
QY 376 GlnValIysAsnValIleTyrGlnAsnIleIysGlyThrSerAlaThrIysValAlaIle 395

```

```

DB 1122 CAATCCAGATGTGTACATACAAAGAACTATCTGGACATTCAGCAACGACAGCAAT 1181
QY 396 LysPheAspCysSerThrAsnThrPheProCysGlnGlyIleIleMetGlnAsnIleAsnLeu 415
DB 1182 CAACCTTAAGTGCAGTACAGTATGCGCTGCAAAAGATATAAGTATGATATCTTGG 1241
QY 416 ValGlyIysSerGlyIysProSerGlnAlaThrCysIysAsnValHisPheAsnVal 434
DB 1242 AAGCTTACCTCAGGAA-----ATTGCTTCCTGCTTAATGATTAATGCAAAATGATAT 1295
QY 435 ---AlaGlnHisValThrProHisCysThrSerLeuGlnIleSer 448
DB 1296 TTCAGTGCACAGTCAATCCCTGATGCAAGAAATTTAATGTCAGT 1340

RESULT 7
US-08-467-023-141
; Sequence 141, Application US/08467023
; Patent No. 6090366
; GENERAL INFORMATION:
; APPLICANT: Griffith, Irwin J.;
; APPLICANT: Pollock, Joanne;
; APPLICANT: Bond, Julian F.;
; APPLICANT: Garman, Richard D;
; APPLICANT: Kuo, Mei-Chang;
; APPLICANT: Yeung, Siu-mei H.;
; APPLICANT: Brauer, Andrew;
; APPLICANT: Exley, Mark A.;
; APPLICANT: Powers, Steven P.
; TITLE OF INVENTION: Allergenic Proteins And Peptides From
; TITLE OF INVENTION: Japanese Cedar Pollen
; NUMBER OF SEQUENCES: 261
; CORRESPONDENCE ADDRESS:
; ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
; STREET: 610 Lincoln St
; CITY: Waltham
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,023
; FILING DATE: June 6, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,225
; FILING DATE: December 6, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 141:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1479 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-467-023-141

Alignment Scores:
Pred. No.: 6,33e-83 Length: 1479
Score: 786.50 Matches: 165
Percent Similarity: 58.66% Conservative: 72
Best Local Similarity: 40.84% Mismatch: 156
Query Match: 32.91% Indels: 11

```

DB: 3 Gaps: 6

US-10-691-374-2 (1-457) x US-08-467-023-141 (1-1479)

QY 53 SerTyriSerLysAsnIleGluSerAspAsnIleAspValAsp----- 69

DB 31 AGTGTGCGAATAATCTTAGATCGAATCGAGTTTAAAGAGTTGAGCTTCTGT 90

QY 70 LysAsnGlyIleValIleAsnValLeuSerPheGlyValAlaGlyAspGlyThr 89

DB 91 CATATGCTATCAACATCTTCAATGCGAAGTAAGTGGCGAGTAGCGATGGAAGCAT 150

QY 90 TyrAspAsnIleAlaPheGluGlnAlaTyrAsnGluAlaCysSerSerArgThrProVal 109

DB 151 GATTGCACTGAGGATTTTCAACAGCATGSCAAGCTGCATCAAAACCATCA---GCA 207

QY 110 GlnPheValValProLysAsnLysAsnTyrLeuLeuGlnIleThrPheSerGlyPro 129

DB 208 ATGTTGCTTGTGCGACGACGAGAAATTTGTTAAACAAATCTTCTTCAATGGGCA 267

QY 130 CysArgSerSerIleSerValLysIlePheGlySerLeuGlnAlaSerSerLysIleSer 149

DB 268 TGTCAACCTCACTTCTTAAAGTAGAGTAGGGAATTAATGCTGCTACCAAAATCCAGCG 327

QY 150 AspTyrLysAspArgArgLeuTyrIleAlaPheAspSerValGlnAsnLeuValValGly 169

DB 328 AGCTGAGAGAAATTAATGATATGTTGAGATTGCTTAAACCTTACAGTTTACTCTATG 387

QY 170 GlyIleGlyThrIleAsnGlyAsnGlyGlnValTyrProSerSerCysLys--Ile 188

DB 388 GGTAAAGGTGTAAATGATGGGCAAGAAACAAATGGTGGCTGCGCAATGAAATGGGTC 447

QY 189 AsnLysSerLeuProCys-----ArgAspAlaProThrAlaLeuThrPheTyrPancCys 206

DB 448 AATGACAGCAAAATTTGCAACAGATCGTATGACCAACAGCATTAATTCATTTTTC 507

QY 207 LysAsnLeuLysValAsnAsnLeuLysSerLysAsnGlnGlnIleHisIleLysPhe 226

DB 508 ACGGCTGTGAATATCCAAAGCATGAACTAATGACATCCCAATTTCAATTAATGTTT 567

QY 227 GluSerCysThrAsnValValAlaSerAsnLeuMetIleAsnAlaSerAlaLysSerPro 246

DB 568 GGGATTTGTGAGGAGTAATAATCATCGCATTAATGATTAATGCGCACGAGACATCTCT 627

QY 247 AsnThrAspGlyValHisValSerAsnThrGlnTyrIleGlnIleSerAspThrIleIle 266

DB 628 AACACTGATGGAATTAATCTTTCATCTTAATAAATCTTCACTTAACAAAGAACAGATA 687

QY 267 GlyThrGlyAspSerCysIleSerIleValSerGlySerGlnAsnValGlnAlaThrAsn 286

DB 688 GGAACAGGGATGATCGCTCGCTAATAGCACAGGGCTTTCTTAATATTTGATGAGAT 747

QY 287 IleThrCysGlyProGlyHisGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlu 306

DB 748 CTGATTTGGGTCGACGATGGAATTAATGAGAGTCTTGGAGGAGAAACCTTGA 807

QY 307 AlaTyrValSerAsnValThrValAsnGlnAlaLysIleGlyAlaGluAsnGlyVal 326

DB 808 GCGAGAGTTTCACTGACGTAATAGGGCTTAATTCATAGACACCAAAATGATTA 867

QY 327 ArgIleLysThrTyrGlnGlyGlySerGlyGlnAlaSerAsnIleLysPheLeuAsnVal 346

DB 868 AGAATCAAAACATGCGAGGTGTTCAAGCATGCGAAGCCATTAATTAATTAATGAGAAATGTT 927

QY 347 GluMetGlnAspValLysTyrProIleIleIleAspGlnAsnTyrCysAspArgValGlu 366

DB 928 GAAATGAATTAATTCGAGAACCCATTAATTAATTAATCACTTCACTCACTTCACTTCT 987

QY 367 ProCysIleGlnGlnPheSerAlaValGlnValLysAsnValValTyrGlnAsnIleLys 386

DB 988 GCTTGCCAAAACAGAGGTCTGCGCTTCAATCCAAAGATGTGACATTAAGAACATATACG 1047

QY 387 GlyThrSerAlaThrLysValAlaIleLysPheAspCysSerThrAsnPheProCysGlu 406

DB 1048 GGGACATCAGCAACAGCAGCATTTCACTTAATGACAGTATGCTGCAAA 1107

QY 407 GlyIleIleMetGluAsnIleAsnLeuValGlyGluSerGlyLysProSerGluAlaThr 426

DB 1108 GATTAATAAGCTAAGTATATATCTTGAAGCTTACCTCAGGAAA-----ATTGCTTCC 1161

QY 427 CysLysAsnValHisPheAsnAsn-----AlaGluHisValThrProHisCysThrSer 444

DB 1162 TGCTTAATGATTAATGCAATGATTAATTCAGTGAACAGTATCTTCATGCAAGAT 1221

QY 445 LeuGluIleSer 448

DB 1222 TTAAGTCCAACT 1233

RESULT 8

US-08-467-023-140

/ Sequence 140, Application US/08467023

/ Patent No. 6090386

/ GENERAL INFORMATION:

/ APPLICANT: Griffith, Irwin J.;

/ APPLICANT: Pollock, Joanne;

/ APPLICANT: Bond, Julian F.;

/ APPLICANT: Garman, Richard D.;

/ APPLICANT: Kuo, Mei-Chang;

/ APPLICANT: Yeung, Siu-mei H.;

/ APPLICANT: Brauer, Andrew;

/ APPLICANT: Ekley, Mark A.;

/ APPLICANT: Powers, Steven P.

/ TITLE OF INVENTION: Allergenic Proteins And Peptides From

/ NUMBER OF INVENTION: Japanese Cedar Pollen

/ CORRESPONDENCE ADDRESSES:

/ ADDRESS: Immunologic Pharmaceutical Corporation, Inc.

/ STREET: 610 Lincoln St

/ CITY: Walham

/ STATE: MA

/ COUNTRY: USA

/ ZIP: 02154

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/467.023

/ FILING DATE: June 6, 1995

/ CLASSIFICATION: 424

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: 08/350,225

/ FILING DATE: December 6, 1994

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Jane E. Remillard

/ REGISTRATION NUMBER: 38,872

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (617) 227-7400

/ TELEFAX: (617) 227-5941

/ INFORMATION FOR SEQ ID NO: 140:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 1395 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: CDNA

/ US-08-467-023-140

Alignment Scores:

Pred. No.: 9.02e-81

Score: 768.00

Percent Similarity: 59.01%

Best Local Similarity: 41.51%

Query Match: 32.13%

Length: 1395

Matches: 159

Conservative: 67

Mismatches: 149

Indels: 8



```

DB:      3      Gaps:      5
US-10-691-374-2 (1-457) x US-08-467-023-140 (1-1395)
QY      71 AenglylleValIleAsnValIleAsnValIleSerPheGlyAlaIleGlySerGlyAspGlyLeuThrTyr 90
      10 GAATGCATCAACACTCTTCAATGTGCAAAAGTATGCGGCACTGAGCATGAAAGCATGAT 69
QY      91 AapAenIleAlaPheGluGlnAlaIlePheGlnAlaIleCysSerSerArgThrProValGln 110
      70 TGCACGTGAGGCAATTTTCAACAGCATGCAAGCTGCAATGCAAAACCACATCA---GCATG 126
QY      111 PheValValProIleAsnValIleAsnValIleLeuLeuLeuGlnIleLeuPheSerGlyProCys 130
      127 TTGCTGTGTCAGGCAAGCAAAATTTGTTGAACATCTGTTCTTCAATGCGGCATG 186
QY      131 ArgSerSerIleSerValIlePheGlySerLeuGlnAlaSerSerIleSerArg 150
      187 CAACCTCATCTTACTTTAAGTAGATGAGATGAGATGATGCTGCTTCAACAAATCCAGGAGC 246
QY      151 TyrIleAspArgArgLeuThrIleAlaPheAspSerValGlnAsnLeuValGlyGly 170
      247 TGGAAATATATGATATGTTGTCAGTTGCTTAACTTACAGGTTTACTTAATGGGT 306
QY      171 GlyGlyThrIleAsnGlyAsnGlyGlnValIleThrProSerSerCysIle---IleAsn 189
      307 AAGGTGTAATGATGCGGCAAGGAAACATGCTGCGTGGCCAAATGTAATGCGTCAT 366
QY      190 LysSerLeuProCys-----ArgAspAlaProThrAlaLeuThrPheThrPheCysIle 207
      367 GAGCGAAGAAATTTGCAACAGATCTGATAGCAACAGCCATTAATTTTCATTTTCCAGC 426
QY      208 AsnLeuLeuValIleAsnLeuLeuSerLysAsnAlaGlnIleIleIleIleLysPheGln 227
      427 GGTCTGATATCCAGAGCTGAATAACTAATGACAGTCCCAATTTTCATTTTGGG 486
QY      228 SerCysThrAsnValIleAlaSerAsnLeuMetIleAsnAlaSerIleLysSerProAsn 247
      487 AATTGAGGAGATTAATAATCATCGCATTTAGTATTACGACCGAGAGACATCTTAC 546
QY      248 ThrAspGlyValIleValSerAsnThrGlnTyrIleGlnIleSerAspThrIleIleGly 267
      547 ACTGATGGAATGATATCTTGTGATCTTAAAACTTTCACTTACAAAGAAACAGATAGCA 606
QY      268 ThrGlyAspAspCysIleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIle 287
      607 ACAGGAGATGACGCTGCTATAGGACACAGGCTTCTTAAATTTGATTTGAGATCTG 666
QY      288 ThrCysGlyProGlyIleGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAla 307
      667 ATTTGGGCTCCAGGCAATGAAATAGATATAGAAAGTCTTGGGAGGAAACTCTTACAGCA 726
QY      308 TyrValSerAsnValIleValIleAsnGlnAlaIleIleGlyIleGlnAsnGlyValArg 327
      727 GAGGTTTCACTACTGACAGTAATGAGGCTTAAATTTCTATAGACACACAAATGATTAAGA 786
QY      328 IleLysThrTyrGlnGlySerGlyGlnAlaSerAsnIleLysPheLeuAsnValGln 347
      787 ATCAAAACATGACGAGGTGGTTCAGGCATGCAAGCCATTAATTTATGAGATGTTGAA 846
QY      348 MetGlnAspValIleTyrProIleIleIleAspGlnAsnTyrCysAspArgValGluPro 367
      847 ATGATTAATTCGAGAAACCCCATATTAATTAATCAATTCACATGCACTTCAGCTTCTGCT 906
QY      368 CysIleGlnIlePheSerAlaValIleValIleAsnValIleTyrGlnAsnIleLysGly 387
      907 TGCCAAAACCAAGAGCTGCGCTTCAATTCAGATGACATGACATCAAGAAACATACGTGG 966
QY      388 ThrSerAlaThrLysValAlaIleLysPheAspCysSerThrAsnPheProCysGlyGly 407
      967 ACATTCAGCAACACAGCAGCAATTCAACTTAAGTGCAGTACAGTATGCTCCCAAGAT 1026
QY      408 IleIleMetGlnAsnIleAsnLeuValIleGlyGlySerGlyLysProSerGlnAlaThrCys 427

```

```

DB      1027 ATAAAGCTAAGATATATCTTTGAAGCTTACTTACAGGAAA-----ATTGCTTCTCC 1080
QY      428 LysAsnValIlePheAsnAsn-----AlaGlnIleValIleProIleCysThrSerLeu 445
      1081 CTTAATGATTAATGCAAAATGATATTTTCACTGACACAGTCATGCTGCATGCAAGAAATTTA 1140
QY      446 GluIleSer 448
      1141 AGTCCAAAGT 1149
DB
RESULT 9
US-08-467-023-139
; Sequence 139, Application US/08467023
; Patent No. 6090386
; GENERAL INFORMATION:
; APPLICANT: Griffith, Irwin J.;
; APPLICANT: Pollack, Joanne;
; APPLICANT: Bond, Julian F.;
; APPLICANT: Garman, Richard D.;
; APPLICANT: Kuo, Wei-Chang;
; APPLICANT: Yeung, Siu-mei H.;
; APPLICANT: Brauer, Andrew;
; APPLICANT: Exley, Mark A.;
; APPLICANT: Powers, Steven P.
; TITLE OR INVENTION: Allergenic Proteins And Peptides From
; TITLE OR INVENTION: Japanese Cedar Pollen
; NUMBER OF SEQUENCES: 261
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Immunologic Pharmaceutical Corporation, Inc.
; STREET: 610 Lincoln St
; CITY: Waltham
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,023
; FILING DATE: June 6, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,225
; FILING DATE: December 6, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1410 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-467-023-139
Alignment Scores:
Pred. No.: 9,18e-81 Length: 1410
Score: 768.00 Matches: 159
Percent Similarity: 59.01% Conservative: 67
Best Local Similarity: 41.51% Mismatches: 149
Query Match: 32.13% Indels: 8
DB: 3 Gaps: 5
US-10-691-374-2 (1-457) x US-08-467-023-139 (1-1410)

```

```

Qy 71 AsnGlyIleuValIleAsnValIleuSerPheGlyValAlaYsglyAspGlyIleThrTy 90
Db 25 GATCTATCAACATCTTCAATGTGAAAGATAGGCGCATGAGCAATGGAAGCATGAT 84
Qy 91 AspAsnIleAlaPheGluGlnAlaIleAsnGlnAlaCysSerSerArgThrProValGln 110
Db 85 TGCACTGAGCATTTTCAACAGCATGCGCAAGCTGATGCAAAACCATCA---GCATG 141
Qy 111 PheValValProValAsnValAsnTyIleuLeuValGlnIleThrPheSerGlyProCys 130
Db 142 TTGCTTGTCCAGGCGCAAGAAATTTGTTAAACATCTGTTCTTCAATGCGGCATGT 201
Qy 131 ArgSerSerIleSerValIlePheGlySerLeuGlnAlaSerSerIleSerArg 150
Db 202 CAACCTCACTTTCTTTAAGTGAATGAGTAAATGCTGCTTACCAAAATCCAGCGGC 261
Qy 151 TyrIleAspArgArgLeuThrIleAlaPheAspSerValGlnAsnLeuValAlaGly 170
Db 262 TGGAAAGATTAATTAATGATATGTTGCAATTTGCTAAACTTACAGTTTACTTATG 321
Qy 171 GlyIleThrIleAsnGlyAsnGlyValIleThrProSerSerCysValys---IleAsn 189
Db 322 AAAGGTGATATGATGGGCAAGAAACATATGCTGGCTGGCCAAATGTAATGGGTCA 381
Qy 190 LysSerLeuProCys-----ArgAspAlaProThrAlaLeuThrPheTrpAsnCys 207
Db 382 GCACGAGAAATTTGCAACGATCGATGATAGACCAAGCATTAATTCATTTTCCACG 441
Qy 208 AsnLeuValAsnAsnLeuValSerValAsnAlaGlnGlnIleHisIleValPheGlu 227
Db 442 GGTCTGATATATCAAGCATGTAATCAATGAAACGATCCGCAATTTCAATTAAGTT 501
Qy 228 SerCysThrAsnValAlaIleSerAsnLeuValIleAsnAlaSerAlaLysSerProAsn 247
Db 502 AATGTGAGGAGGAGTAAATCATCGCATTAATGATTAAGGACCGCAAGACATCTTAC 561
Qy 248 ThrAspGlyValIleValSerAsnThrGlnIleGlnIleSerAspThrIleIleGly 267
Db 562 ACTGATGTAATGATATCTTTCGATCTTAAACCTTTCACTTCAAAAAGAACACAGAT 621
Qy 268 ThrGlyAspArgCysIleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIle 287
Db 622 ACAGGGATGATGCTGCTGCTATGAGCAAGGGCTTCTTAATTTGTGATTGAGATCTG 681
Qy 288 ThrCysGlyProGlyIleGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAla 307
Db 682 ATTTGCGGTCAAGCGCATGGAATTAATGATTAAGTCTTGGGAGGAAACTCTAGAGA 741
Qy 308 TyrValSerAsnValThrValAsnGlnAlaValIleIleGlyValGlnAsnGlyValArg 327
Db 742 GAGGTTTCATACGTGACGTAATGAGGCTTAATTCATAGACACACAAATGATTAAGA 801
Qy 328 IleIleThrThrPheGlnIleGlySerGlyGlnAlaSerAsnIleValPheLeuAsnValGlu 347
Db 802 ATCAAAACATGCGAGGGGTGTTCAAGCATGCGCAAGCATATTAATTAAGAAATTTGA 861
Qy 348 MetGlnAspValIleTyProIleIleIleAspGlnAsnTyCysAspArgValGluPro 367
Db 862 ATGATTAATTCGAGAAACCCATATTAATTAATCAATTTCTACTGACTTCAGCTTGT 921
Qy 368 CysIleGlnIlePheSerAlaValGlnValIleValAsnValIleTyGlnAsnIleGly 387
Db 922 TGCCAAACATGAGGGGTGCTGCGTTCAAAATCCAAATGTCATCAACAGACATACGTGG 981
Qy 388 ThrSerAlaThrIleValAlaIleIleValPheAspCysSerThrAsnAspPheProCysGlu 407
Db 982 ACATCAGCAACAGCAGCATTTCACTTAAGTCAGTACAGTACAGTACCTGCAAGAT 1041
Qy 408 IleIleMetGlnAsnIleAsnLeuValIleGlySerGlyIleProSerGlnAlaThrCys 427
Db 1042 ATTAAGCTTAAGTATTAATTTTGAAGCTTACTTCAAGGAAA-----ATTGCTTCTG 1095
Qy 428 LysAsnValIleIlePheAsnAsn-----AlaGlnIleIleValIleThrProHisCysThrSerLeu 445

```

```

Db 1096 CTTATATATATATGCAATGATATTTTCAATGAGACAGCATCTCCGATGCAAGAAATTTA 1155
Qy 446 GluIleSer 448
Db 1156 AGTCCAAAGT 1164

RESULT 10
US-08-463-213-1
Sequence 1, Application US/08463213
Patent No. 5759829
GENERAL INFORMATION:
APPLICANT: SHEWMAKER, C.
APPLICANT: KRIDL, J.
APPLICANT: HIATT, W.
APPLICANT: KNAUF, V.
TITLE OF INVENTION: ANTI-SENSE REGULATION OF GENE
TITLE OF INVENTION: EXPRESSION IN PLANT CELLS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSER: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1 (a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,213
FILING DATE: 5-JUNE-95
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/750,505
FILING DATE: 27-AUG-91
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/240,408
FILING DATE: 30-AUG-88
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/920,574
FILING DATE: 17-OCT-86
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/845,676
FILING DATE: 28-MAR-86
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Laesen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 26-4
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 219 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-463-213-1

Alignment Scores:
Pred. No.: 2,15e-34 Length: 219
Score: 366.00 Matches: 70
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.59% Mismatches: 0
Query Match: 15.31% Indels: 0

```

```
DB: 1 Gaps: 0
US-10-691-374-2 (1-457) x US-08-463-213-1 (1-219)
QY 71 AaaglyllylVallleasvallyleuSerPhegllyAlalyeglyVaspglylysthrTy 90
    7 CATGGGATTAAAGTGAATTATGACTTACCTTTGGAGCTAAGGCTGATGAAAAACATAT 66
QY 91 AaPaenllleAlaPhegluGlnAlaTTPaangluAlaCyseSerSerArgThrProValGln 110
    67 GATTAATATGCAATTGAGCAAGCATGGAATGAAGCATGTTCATCTGAAACACCTGTTCAA 126
QY 111 PheValValProlysaenlysaenTyrlleuleuLyGlnlleThrpheSerGlyProCy 130
    127 TTGTGGTTCCTAAACAAAGAAATTATCTTCACAGCAATACCTTTTCAGGTCATCC 186
QY 131 ArgSerSerlleSerVallylsilePhegllySer 141
    187 AGATCTTCTATTTCAGTAAAGATTTTGGATCC 219
DB 187 AGATCTTCTATTTCAGTAAAGATTTTGGATCC 219

RESULT 11
5453566-1
; Patent No. 5453566
; APPLICANT: SHEWMAKER, CHRISTINE K.; KRIDL, JEAN C.; HIATT,
; WILLIAM R.; KNAUF, VIC
; TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION
; IN PLANT/CELLS
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/750,505
; FILING DATE: 27-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 240,408
; FILING DATE: 30-AUG-1988
; APPLICATION NUMBER: 920,574
; FILING DATE: 17-OCT-1986
; APPLICATION NUMBER: 845,676
; FILING DATE: 28-MAR-1986
; SEQ ID NO:1:
; LENGTH: 219
5453566-1

Alignment Scores:
Pred. No.: 2,15e-34 Length: 219
Score: 366.00 Matches: 70
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.59% Mismatches: 0
Query Match: 15.31% Indels: 0
Gaps: 0
DB: 6 Gaps: 0

US-10-691-374-2 (1-457) x 5453566-1 (1-219)
QY 71 AaaglyllylVallleasvallyleuSerPhegllyAlalyeglyVaspglylysthrTy 90
    7 CATGGGATTAAAGTGAATTATGACTTACCTTTGGAGCTAAGGCTGATGAAAAACATAT 66
QY 91 AaPaenllleAlaPhegluGlnAlaTTPaangluAlaCyseSerSerArgThrProValGln 110
    67 GATTAATATGCAATTGAGCAAGCATGGAATGAAGCATGTTCATCTGAAACACCTGTTCAA 126
QY 111 PheValValProlysaenlysaenTyrlleuleuLyGlnlleThrpheSerGlyProCy 130
    127 TTGTGGTTCCTAAACAAAGAAATTATCTTCACAGCAATACCTTTTCAGGTCATCC 186
QY 131 ArgSerSerlleSerVallylsilePhegllySer 141
    187 AGATCTTCTATTTCAGTAAAGATTTTGGATCC 219
DB 187 AGATCTTCTATTTCAGTAAAGATTTTGGATCC 219

RESULT 12
5453566-1
; Patent No. 5453566
; APPLICANT: SHEWMAKER, CHRISTINE K.; KRIDL, JEAN C.; HIATT,
; WILLIAM R.; KNAUF, VIC
```

```
; TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION
; IN PLANT/CELLS
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/750,505
; FILING DATE: 27-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 240,408
; FILING DATE: 30-AUG-1988
; APPLICATION NUMBER: 920,574
; FILING DATE: 17-OCT-1986
; APPLICATION NUMBER: 845,676
; FILING DATE: 28-MAR-1986
; SEQ ID NO:1:
; LENGTH: 219
5453566-1

Alignment Scores:
Pred. No.: 2,15e-34 Length: 219
Score: 366.00 Matches: 70
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.59% Mismatches: 0
Query Match: 15.31% Indels: 0
Gaps: 0
DB: 6 Gaps: 0

US-10-691-374-2 (1-457) x 5453566-1 (1-219)
QY 71 AaaglyllylVallleasvallyleuSerPhegllyAlalyeglyVaspglylysthrTy 90
    7 CATGGGATTAAAGTGAATTATGACTTACCTTTGGAGCTAAGGCTGATGAAAAACATAT 66
QY 91 AaPaenllleAlaPhegluGlnAlaTTPaangluAlaCyseSerSerArgThrProValGln 110
    67 GATTAATATGCAATTGAGCAAGCATGGAATGAAGCATGTTCATCTGAAACACCTGTTCAA 126
QY 111 PheValValProlysaenlysaenTyrlleuleuLyGlnlleThrpheSerGlyProCy 130
    127 TTGTGGTTCCTAAACAAAGAAATTATCTTCACAGCAATACCTTTTCAGGTCATCC 186
QY 131 ArgSerSerlleSerVallylsilePhegllySer 141
    187 AGATCTTCTATTTCAGTAAAGATTTTGGATCC 219
DB 187 AGATCTTCTATTTCAGTAAAGATTTTGGATCC 219

RESULT 13
US-09-107-532A-2998
; Sequence 2998, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
```

NAME: Arinello, Pamela Deneke  
 REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: GTC-012  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781) 893-5007  
 TELEFAX: (781) 893-8277  
 INFORMATION FOR SEQ ID NO: 2998:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1356 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Enterococcus faecium  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (B) LOCATION 1...1356  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2998:  
 US-09-107-532A-2998

Alignment Scores:  
 Pred. No.: 4,06e-25 Length: 1356  
 Score: 299.00 Matches: 100  
 Percent Similarity: 45.98% Conservative: 66  
 Best Local Similarity: 27.70% Mismatches: 134  
 Query Match: 12.51% Indels: 62  
 DB: 4 Gaps: 15

US-10-691-374-2 (1-457) x US-09-107-532A-2998 (1-1356)

QY 73 l l e l y s v a l l e a s n v a l l e u s e r p h e g l y a l a l y s g l y a s p g l y s t h r t y r a s p a n 92  
 D b 13 a t t a t a t a t g a t g a t a t t t t g a c a t t t g c g c a c a t t g a t -----g a a t t g a a t a c t 66  
 QY 93 l l e l a p h e g l u g l i n a l t r p a n g l u a l c y s e s e r a r g t h r p r o v a l g l n p h e a l 112  
 D b 67 g a n c a n t a c a a c c a t t g a t g c t g c a g c a a g t a c g a g c g a g a c c ---g r a g t t 123  
 QY 113 v a l p r o l y s a n l y s a n t y r l e u l e u l y e g l n l e t h r p h e s e r g l y p r o c y s a g s e r 132  
 D b 124 g t t c a c a g a g t g a a ---t t t t a c a g a g c t t a t t t t t a a a a c a a t g t c g a a c t 179  
 QY 133 s e r i l e s e r v a l l y s i l e p h e g l y s e r l e u g l u a l a s e r s e r l y s i l e s e r a s p t y r l y 152  
 D b 180 t c a t t t a a g c g a g a ---g c g g t t c g a a a t t c t g a c a t c c a a a a g a t t a t t a c c 233  
 QY 152 s -----a s p a r g a r g l e u t r p -----l l e a l a p h 160  
 D b 234 t g t t g t t a t t c a c g t t g g a g a g a t c c a t c g a a a a g t t t a t g c t t c c t g a t t a t g c 293  
 QY 160 e a a p s e r v a l g l n a n l e u v a l a l y g l y g l y t h r l l e a n g l y a n g l y g l n a 180  
 D b 294 t c a a a a t g a g a a c a t c t g t c a c t g t t c g t a c a t a g a t g a a n t g a a a a a 353  
 QY 180 l t p t r p -----p r o s e r s e r c y s l s i l e a n l y s e r l e u p r o c y 194  
 D b 354 a t g t g c a t a c t t t t a g a a a t g a a c a c a t -----a a t t c g c t t a t c c t - 399  
 QY 194 s a r g a s p a l a p r o t h r a l a l e u t h r p h e t r p a n c y s l y s a n l e u l y s v a l a n a n l e 214  
 D b 400 -----c g t c c t a a t t g a t a g t t c a t a t g t c a t g a t t a c t g a a a g a t a t 452  
 QY 214 u l y s e r l y s a n a l a g l n g l n l l e h i s l e l y s p h e g l u s e r c y s t h r a n v a l a l 234  
 D b 453 t a a a t t g a t c c a c a c a a g t t g a c a t a c c c a a t t c t t t g c a g t a t a c a c g t t 512  
 QY 234 a s e r l e u w e t l l e a n a l a s e r a l a l y s e r p r o a n t h r a s p g l y a l h i s v a l s e 254  
 D b 513 t g a t a c t t t g a c a t t t g a a c c a g c g a t t a c c a a a t t a c a g a a t t a t t a t c a a 572

QY 254 r a s n t h r g l n t y r l l e g l n l l e s e r a s p t h r l l e i l e g l y t h r g l y a s p a s c y e i l s e 274  
 D b 573 a t c t a t g c a a a a a a t g c g a t t a t t a g t a t t g a c a t a t t g a c g t g a g a t a c t g a t t g c 632  
 QY 274 r l l e a l s e r g l y s e r -----g l n a n v a l g l n a l e t h 285  
 D b 633 a a t c a a g c a g a c t g a a n a c g a a n a c g a a a g a t c c t t c g a a a a a t a t c a c a t a c 692  
 QY 285 r a s n l l e t h r c y e g l y p r o g l y h i s ---g l y l l e s e r l l e g l y s e r l e u g l y s e r g l y a s 304  
 D b 693 a a a t t c a c a g a t g c a t g a c a t g c g c g t g c g t a t t a g a a g a t g a a n a g t g t a g 752  
 QY 304 n s e r g l u a l a t y r v a l s e r s a n v a l t h r v a l a n g l u a l a l y s l l e i l e g l y a l a g l u a s 324  
 D b 753 t -----a t t c g a a a t a t t a c a t t t c a a t t g a t t c t t c c a a a a a c t g a t c g 800  
 QY 324 n g l y a l a g l l e l y s t h r t p g l n g l y g l y s e r g l y c l n a l a s e r a n l l e l y s p h e l e 344  
 D b 801 a g a a t a c a t t g a a a t c t a g a c a g t c g a g c g g a t c g t t g a a g a t a t t c a g a t a t c 860  
 QY 344 u a s n v a l g l u w e t g l n a s p v a l y s t y r p r o l l e i l e l e a s p g l n a s n t y r ---c y s a s 363  
 D b 861 c a a t a t a c t g a t a t a t g a t g a t g c t c a t t t a t a c t g a a t c t t a t t a t t a t t t c g t g 920  
 QY 363 p -----a r g v a l g l u p r o c y s i l e g l n g l n ----- 371  
 D b 921 t c c g c g a g a a a a c c g t a t g t a t g g a a a a a a a a c t t a t c c a t a g a t a c a c g a c 980  
 QY 372 -----p h e s e r a l a l a g l n v a l l y s a n v a l a t y r g l u a s n l l e l y s g l y t h r s e 389  
 D b 981 a c c a g a t t c a g a a g a t c a t t t t t g a a t a t c a c g c a a g a a a t g t c a t c t t c t g c 1040  
 QY 389 r 389  
 D b 1041 A 1041  
 RESULT 14  
 US-09-614-221A-573  
 Sequence 573, Application US/09614221A  
 Patent No. 6723837  
 GENERAL INFORMATION:  
 APPLICANT: Karunanandaa, Balasubramanian  
 APPLICANT: Yu, Jaehyuk  
 APPLICANT: Kishore, Ganesh M.  
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED  
 TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM  
 FILE REFERENCE: 16516.075  
 CURRENT APPLICATION NUMBER: US/09/614,221A  
 PRIOR FILING DATE: 2000-07-12  
 PRIOR APPLICATION NUMBER: US 60/142,981  
 NUMBER OF SEQ ID NOS: 626  
 SEQ ID NO 573  
 LENGTH: 1086  
 TYPE: DNA  
 ORGANISM: Saccharomyces cerevisiae  
 US-09-614-221A-573

Alignment Scores:  
 Pred. No.: 6,37e-25 Length: 1086  
 Score: 296.00 Matches: 112  
 Percent Similarity: 44.85% Conservative: 58  
 Best Local Similarity: 29.55% Mismatches: 147  
 Query Match: 12.38% Indels: 62  
 DB: 4 Gaps: 17

US-10-691-374-2 (1-457) x US-09-614-221A-573 (1-1086)

QY 61 s e r a n s a n a n l l e a s p l y s a n g l y l l e l y s v a l l e a s n v a l l e u s e r 80  
 D b 97 t c t t g t c t c a c t c t c a c c g t g a a a a a t t g n a c a g a c t c g t a t t a a a g c t t a a c t 156  
 QY 81 p h e g l y a l a l y s g l y a s p g l y s t h r t y r a s p a n l l e a l a p h e g l u g l n a l a t r p a n 100

Db	157	GTCCACGT-----GACAGACCTTAAGTTTAACGGGTTA-----	192
Oy	101	GIUALAcySeSerSerargThrpProValGlnpheValProLyAsnLysAntThyreU	120
Db	193	-----AGCAGGTGTACTACTGTTACGTTTGAAGGACAAACACATTTCAGTAACAAG	243
Oy	121	LeuYvGlnIleThrPheSerGlyProCyAsArgSerIleSerValLysIlePheGly	140
Db	244	---GAATGGAGCGGCCCTTTA-----ATTTCATC-----	270
Oy	141	SetLeuGIUALaSerSerLysIleSerAerTyLysAspArgLentrpIleAlaPhe	160
Db	271	-----TCGGGCTTAATAATCAGC-----	288
Oy	161	AspSerValGlnAsnLeuValGlyGlyGly---ThrIleAsnGlyAsnGlyGln	179
Db	289	-----GTTTGTGGTGTGGGACATACCATTTGATGTCAAGAGACA	330
Oy	180	ValTrpTrpProSerSerCyLysLysIleAsnLysSerLeuProCyAsArgIleAspIleProthr	199
Db	331	AAATGGTGGAGTGGCTTAGTGATGACGGTAAAGTCMAACCG---AAGTTGTAAAGTTG	387
Oy	200	AlaLeuThrThrPheTrpAsnCyLysAsnLeuLysValAsnAsnLeuLysSerLysAsnIle	219
Db	388	CGGTGAGC-----GGAACATCTAAGTCACCGAGTTGAATTTTAAAAATGT	435
Oy	220	GlnGlnIleAsIleLysPheGluSerCyThrAsnValValAserAsnLeuMetIle	239
Db	436	CCACACCAAGTTCACGATCACTAAATGTTACAGTTTAAACATCAGCACATTAACATT	495
Oy	240	AsnAla-----SerAlaLysSerProAsnThrAspGlyValHisValSerAsn	255
Db	496	GATATCAGAGACGCTGATCGGCTGGTGTCTATAATACGATGGGTTGATGTGGTAT	555
Oy	256	ThrgIntTyrlleGlnIleSerAspThrIleIleGlyThnGlyAspAspCysIleSerIle	275
Db	556	TCTAGTAACTTTAATTCAAGAGATGACTGTTTAAATCAGAGATGACGTATTCTGTG	615
Oy	276	ValSerGlySerGlnAsnValGlnAlaThrAsnIleThnCyGlyProGlyIleGlyIle	295
Db	616	AATTCGGTTC---ACTATTAAATTATGAAACAACACTGCTTACATAGCCATGATATT	672
Oy	296	SerIleGlySerLeuGlySerGlyAsnSerGlnAlaTyValSerAsnValThrValAsn	315
Db	673	TCTGAGGTTCTGTTGGT---GGCCGTTCTGATAATACAGTCAATGGTTCTGGGGTGA	729
Oy	316	GIUALaLysIleIleGlyAlaGluAsnGlyValAlaGlyIleLysThrTrpGlnGlyGlySer	335
Db	730	AATAACAATGTTATCACTGACACACGGGTGAGAAATMAAACCGTIGAAAGTGCACACA	789
Oy	336	GlyGlnAlaSerAsnIleLysPheLeuAsnValGluMetGlnAspValLys---TyrPro	354
Db	790	GGCAGACGTCACTAAATGTCACTTTATACGTATAATAAATTAAGCGGCATTAATAATGAT	849
Oy	355	IleIleIleAspGlnAsnTyrcyAsAspArgValGluProCyIleGlnGlnPheSerAla	374
Db	850	ATTGTTATCGAAGGGATTTATTTGAATAGT---AAGACTACTGGAACtGCTACAGTGGC	906
Oy	375	ValGlnValLysAsnValValTyrgluAsnIleLysGly-----ThrSerAlaThr	391
Db	907	GTTCCCATTTCCGAATTTAGTATGAAGGATATCACCGGGAGCGTGAATCTCCACACAGAG	966
Oy	392	LysValAlaIleLysPheAspCysSerThrAsnPheProCyGlnGlyIle---IleMet	410
Db	967	AGGGTTAAATTTGGTGAAAAACCTATTAATCGCAATGGCTGGGGGTGCAATTAC	1022
Oy	411	GluAsnIleAsnLeuValGlyGluSerGlyLysPro-----SerGluAlaThrCys	427
Db	1027	GGTGTTCTTCTATTCTGAGAGTTCTGGAATCCCATCTGATCTGCTGCAAGCTGT	1083

```

: Sequence 1, Application US/09787583
: Patent No. 6502696
: GENERAL INFORMATION:
: APPLICANT: DSM N.V.
: APPLICANT: Franse, Maartje
: APPLICANT: Graesslin, Catherine
: APPLICANT: Herweijer, Margareta
: APPLICANT: Meeuwse, Petrus
: APPLICANT: Ooljen, Albert
: APPLICANT: Voragen, Alphons
: TITLE OF INVENTION: ASPERGILLUS TUBIGENSIS POLYGALACTURONASE
: FILE REFERENCE: 24615-2014.00
: CURRENT APPLICATION NUMBER: US/09/787, 583
: PRIOR FILING DATE: 2001-03-19
: PRIOR APPLICATION NUMBER: EP 99200491.2
: PRIOR FILING DATE: 1999-02-17
: PRIOR APPLICATION NUMBER: EP 98203171.8
: PRIOR FILING DATE: 1998-09-18
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 1570
: TYPE: DNA
: ORGANISM: Aspergillus tubingensis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (172) ... (1260)
US-09-787-583-1

Alignment Scores:
Score: 2,33e-24 Length: 1570
Pct: 293.50 Matches: 81
Percent Similarity: 50.54% Conservative: 59
Best Local Similarity: 29.24% Mismatches: 96
Query Match: 12.28% Indels: 41
DB: Gaps: 14

US-10-691-374-2 (1-457) x US-09-787-583-1 (1-1570)
QY 171 GYGYLYTHRIIEASNGLYASNGLYGLNVALTPTP----- 162
DB 475 GGTGCGAAGATCACTCGACGCTGCTCGCTGTGGAGCGGCAAGGTAGCAACGGTGC 53
QY 183 -----ProSerSerCysLysLysLeuLysSerLeuProCysAlaAlaProThr 199
DB 535 AAGACCAAGCCCAAGTTCTTCAGGTCACAG----- 567
QY 200 AlaleuthrPheTrpAsnCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsnAla 219
DB 568 ---CTCGACGAGTCCGACATCCCGGCTGGAAGATCTACAACACCCCTGCGGGCTTC 624
QY 220 GINGNILE-----HisIleLysPheGluSerCysThrAsnValAlaLaserAsnLeu 237
DB 625 AGCATTTGGCTGACCACTG-----ACCATCATGACGTGACCATTAGCAACTCC 675
QY 238 MetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGln 257
DB 676 GCC-----GGACGACGACAGGCGCCACAACACGATGCTTTGACATTGCTCAGAGTACC 729
QY 258 TyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSer 277
DB 730 TACATCACCATGACGAGTGGCCACCGCTCTCAACACAGATGATGCTGCGCATCACTCG 769
QY 278 GYSerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIle 297
DB 790 GGT---GAGCAATCATCTTTCACCAACGGTTACTGTGACGAGTGGCCACGATCTCTCCATT 846
QY 298 GYSerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAla 317
DB 847 GGTTCATTGTTGTT---GGCCGACGACACAACACCTGCAACGAGCTGACATCTCCACTCC 903
QY 318 LysIleIleIleGlyAlaGlnAsnGlyValAlaArgIleLysThrTyrPheGlnGlySerGlyGln 337

```

```
Db      904 AAGTGCTCACTCCGAAACGGTGTCCGTATCAGACCATCTACGGCAAGACCGGCACT 963
Qy      338 AlaSerAsn11elysPheLeuAsnVal1GluMetGlnAspVal---LysTyrPro11e1le 356
Db      964 GTTGAGAACGTCMAAGTTGAGAGCATCATCCCTGTCCGACATCAGCAAGTACGGTATCGTC 1023
Qy      357 11eAspGlnAsnTyrCyAspArgVal1GluProCys11eGlnGlnPheSerAlaVal1Gln 376
Db      1024 GTTGAGCAGGACTAC---GAGAACGGCAGCCGCCACCGGCACGCCCAACCAACGGTGTCAAG 1080
Qy      377 Val1LysAsnVal1Val1TyrGlnAsn11elysGly-----ThrSerAlaThrLysVal 393
Db      1081 GTTGAGGACATCATCTTCAAGAGGTCAACCGGACCGTCACAGAGCTGTGTACTGACATC 1140
Qy      394 Ala11elysPheAspCyserThrAsnPheProCyGluGly11e11eMetGluAsn11e 413
Db      1141 TACATCTCTG-----TGGGTTCCGGC---AGCTGCTCGAACTGGACCTGGAGCGGTGTT 1191
Qy      414 AsnLeuVal1GlyGluSerGlyLysProSerGluAlaThrCysLysAsnVal 430
Db      1192 GATGTGACCGGC-----GGCAAGAAAGAC---AGCAAGTGCAGAAACGTC 1233
```

Search completed: March 5, 2005, 02:29:24  
Job time : 251 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM proteoin - nucleic search, using frame\_plus\_p2n model

Run on: March 4, 2005, 12:45:39 ; Search time 640 Seconds  
(without alignments)  
4227.065 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MTIGNSILLIIIFASIS.....VTPHCTSLISDEALVNY 457

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xlp  
-Q/cg2.1/USPTO.epool.p/US10691374/runat.28022005.120705.20976/app\_query.fasta\_1.647  
-DB=N Geneseq.16Dec04 -QFMT=fastap -SUFFIX=mg -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNIT=bites -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptco -NORExt -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10691374.@CGN\_1\_1\_708.@runat.28022005.120705.20976 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq.16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2390	100.0	1634	1	AAN91112
2	2390	100.0	1636	2	AAQ38415
3	2390	100.0	1636	4	AAQ38415
4	2390	100.0	1637	1	AAN80487
5	1820	76.2	5832	6	AAS20853

6	1127	47.2	1296	6	ADG87885
7	1127	47.2	1296	3	ADA68531
8	1127	47.2	1296	3	AAQ6786
9	1115.5	46.7	1280	8	ADA68538
10	1113	46.6	1631	2	AAT63603
11	1103	46.2	1657	2	AAT63394
12	1103	46.2	1657	2	AAZ22980
13	1022.5	42.8	1659	8	ADA70539
14	1011	42.3	1182	8	ADA70540
15	999	41.8	1344	8	ADA70633
16	971.5	40.6	1168	12	ADJ39618
17	925	38.7	1603	4	AAQ03721
18	856.5	35.8	4783	2	AAQ56765
19	845	35.4	1505	3	AAQ50995
20	841.5	35.2	1236	8	ADA69348
21	829	34.7	5002	2	AAQ25011
22	791.5	33.1	1772	2	AAT38521
23	787.5	32.9	1545	2	AAQ90156
24	787	32.9	1542	2	AAQ84045
25	787	32.9	1726	2	AAQ66048
26	787	32.9	1733	2	AAQ84046
27	787	32.9	1733	2	AAT18102
28	786.5	32.9	1479	2	AAQ66051
29	768	32.1	1395	2	AAQ66050
30	768	32.1	1410	2	AAQ66049
31	767	32.1	1380	2	AAQ84044
32	760	31.8	1524	4	AAI66167
33	760	31.8	1759	4	AAI66166
34	724	30.7	1512	8	ADA70810
35	727	30.4	1673	6	AAQ32961
36	718	30.0	1492	3	AAQ32994
37	709.5	29.7	1501	3	AAQ46834
38	702.5	29.4	1185	6	ABZ14715
39	694	29.1	1155	8	ADA67805
40	694	29.0	1496	2	AAQ56767
41	690	28.9	1736	3	AAQ42340
42	676.5	28.3	1531	3	AAQ49688
43	673	28.2	1690	3	AAQ39627
44	658.5	27.6	1059	3	AAQ36613
45	652	27.3	1182	8	ADA70913

#### ALIGNMENTS

RESULT 1	
ID	AAN91112 standard; DNA; 1624 BP.
XX	
AC	AAN91112;
XX	
DT	25-MAR-2003 (revised)
DT	21-JUN-1990 (first entry)
XX	
DE	Polygalacturonase (PG) gene cDNA.
XX	
KW	Polygalacturonase gene; tomato; ds.
XX	
OS	Lycopersicon esculentum.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
XX	47..1418
XX	/tag= a
XX	
XX	US4801540-A.
XX	
XX	31-JAN-1989.
XX	
PF	02-JAN-1987; 87US-00000201.
XX	
XX	28-MAR-1986; 86US-00845676.
PR	17-OCT-1986; 86US-00920574.
XX	
PA	(CALF) CALGENE INC.

Adg87885	A. thaliana
ADA68531	Arabidops
AAQ6786	Arabidops
ADA68538	Arabidops
AAT63603	Dehiscenc
AAT33994	Olibeed r
AAZ22980	Brassic
ADA70539	Rice gene
ADA70540	Rice gene
ADA70633	Rice gene
ADJ39618	Plant CDN
AAQ03721	Tomato po
AAQ56765	Putative
AAQ50995	Arabidops
ADA69348	Rice gene
AAQ25011	Anther-ep
AAT38521	Chamaecyp
AAQ90156	Japanese
AAQ84045	Japanicum
AAQ66048	Japanese
AAQ84046	Japanicum
AAT18102	Japan ced
AAQ66051	Japanese
AAQ66050	Japanese
AAQ66049	Japanese
AAQ84044	Japanicum
AAI66167	Unipertus
AAI66166	Unipertus
ADA70810	Rice gene
AAQ32961	Pear poly
AAQ32994	Arabidops
AAQ46834	Arabidops
ABZ14715	Arabidops
ADA67805	Arabidops
AAQ56767	Putative
AAQ42340	Arabidops
AAQ49688	Arabidops
AAQ39627	Arabidops
AAQ36613	Arabidops
ADA70913	Rice gene

XX Hiatt WR, Sheehy RE, Shewmaker CK, Kridl JC, Knauf V,  
 XX WPI: 1989-053640/07.  
 DR P-PSDB; AAP94619.  
 XX  
 PT Tomato polygalacturonase gene - used for modulating expression in plant  
 cells or directing expression of heterologous peptide(s).  
 XX  
 PS Disclosure; Page ?; bpp; English.  
 XX  
 CC cDNA of gene may act as a probe to the genomic sequence, provides a means  
 CC of modulating the production of PG and acts as a source of the transit  
 CC peptide which may be joined to heterologous peptides directing them to  
 CC cell wall. (Updated on 25-MAR-2003 to correct PA field.)  
 XX

SQ Sequence 1624 BP; 586 A; 231 C; 294 G; 513 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	6,03e-219	Length:	1624
Score:	2390.00	Matches:	457
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Gaps:	0
DB:	1		

US-10-691-374-2 (1-457) x AAN91112 (1-1624)

QY 1 MetValIleGlnArgAsnSerIleuLeuLeuIleIleIlePheAlaSerSerIleSer 20  
 Db 47 ATGGTATCCAAAGGAATAGTAATCTCTCTCATATATATTTTGGTTCATCAATTTCA 106  
 QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheGlyGlnValTyrAspAsnIleLeu 40  
 Db 107 ACTGTGACAAACATGATATGATGACATTAATTAACAAGTTATGATATATTTCTT 166  
 QY 41 GlnGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerTyrAsnIleGlu 60  
 Db 167 GAACAAAGATTTGCTCATGATTTTCAAGCTTATCTTATTTGAGCAAAATATATGAA 226  
 QY 61 SerAsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValIleAsn 80  
 Db 227 AGCAACATATATATTTGACAAAGTTGATTAATGGAATTAATGATTAATGATTAATG 286  
 QY 81 PheGlyAlaLysGlyAspGlyLysGlyTyrAspAsnIleAlaPheGlnGlnAlaTyrAsn 100  
 Db 287 TTGGAGCTAAGGATGATGAAAAACATATGATATATGCAATTTGAGCAAGCATGAT 346  
 QY 101 GlnAlaCysSerSerArgThrProValGlnPheValValProLysAsnLysAsnTyrLeu 120  
 Db 347 GAAGCATGTTCACTGAAACACCTGTTCAATTTTGCTTCTTAATAAACAAGATTAATCTT 406  
 QY 121 LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly 140  
 Db 407 CTCACCAATATCACTTTTCAGGTCATGACATCTTCAATTCAGTAAGAATTTTGA 466  
 QY 141 SerLeuGlnAlaSerSerLysIleSerAspTyrLysAspArgLysLeuTyrIleAlaPhe 160  
 Db 467 TCCTTGAAAGCATCTGTAATAATTCAGACTCAAAAGATGAAGGCTTTGGATGCTTTT 526  
 QY 161 AspSerValGlnAsnLeuValValGlyGlyGlyTyrIleAsnGlyAsnGlyGlnVal 180  
 Db 527 GATAGGTTCAAAATTAATTTAGTTGTGAGAGAGAGAACTATCAATGAGCAAGCAAGTA 586  
 QY 181 TrpTyrProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaProThrAla 200  
 Db 587 TGTGTGCCAAGTTTCAGCAAAATTAATTAATCACTGCATGACAGGATGACCAACGCGCC 646  
 QY 201 LeuThrPheTyrAsnCysLysAsnLeuLysValAsnAsnLeuLysSerTyrAsnAlaGln 220  
 Db 647 TTAACTCTTGGAATTTGCAAAATTTGAAAGTGAATTAATCTAAAGAGTAAGTAATGCAAA 706  
 QY 221 GlnIleHisIleLysPheGluSerCysThrAsnValValAlaSerAsnLeuMetIleAsn 240

Db 707 CAATTATCATCAAAATTTGATGATCATCAATATGTTGAGTCTCAAAATTTGATGATCAAT 766  
 QY 241 AlAserAlaLysSerProAsnThrAspGlyValIleValSerAsnThrGlnTyrIleGln 260  
 Db 767 GCTTCAGCAAAAGGCCCAATCTGATGAGTCCATGATCAAAATATCAATATATTGCA 826  
 QY 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280  
 Db 827 ATATCTGATACATTTATTTGAAACAGGTGATGATGATTAATTTCAATTTGTTCTGGATCTCA 886  
 QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyIleGlyIleSerIleGlySerLeu 300  
 Db 887 AATGTCAGAGCCCAAAATATTAATCTTGTTGAGTCCAGTCAATGATTAAGTATGGAAGCTTA 946  
 QY 301 GlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaLysIleIle 320  
 Db 947 GGAATCTGAAATTCAGAAAGCTTATGTCATATGTTACTGTAAATGAAGCCAAATATATC 1006  
 QY 321 GlnAlaGlnAsnGlyValArgIleLysThrTyrGlnGlyGlySerGlyGlnAlaSerAsn 340  
 Db 1007 GGTGCCGAAATGAGGTTAGATCAAGCTTGCAAGGAGGATCTGCAACAGCTTAGCAAC 1066  
 QY 341 IleLysPheLeuAsnValGlnMetGlnAspValLysTyrProIleIleIleAspGlnAsn 360  
 Db 1067 ATCAAAATTTCTGAATGTGAAATGCAAGACCTTAAGTATCCATTAATATGACCAAAAC 1126  
 QY 361 TyrCysAspArgValGlnProCysIleGlnGlnPheSerAlaValGlnValLysAsnVal 380  
 Db 1127 TATGTGATGCAAGTTGAACCATGATACACAGTTTTCAGAGTTCAAGTGAAGAAAGTG 1186  
 QY 381 ValTyrGlnAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAspCysSer 400  
 Db 1187 GTGATGGAATATCAAGGCGACAAAGTCAAAAGGTGCCATAAATTTGATTGGAC 1246  
 QY 401 ThrAsnPheProCysGlnGlyIleIleMetGlnAsnIleAsnLeuValGlyGlySerGly 420  
 Db 1247 ACAAACTTTCATGATGAAGAAATTAATATGAGAAATTAATTAATTAATTAATTAAT 1306  
 QY 421 LysProSerGlnAlaThrCysLysAsnValHisPheAsnAsnAlaGlnHisValThrPro 440  
 Db 1307 AAACCATCAGAGGCTAGTCAAAATATGTCATTTTAACAATCTGAACATGTTACACA 1366  
 QY 441 HisCysThrSerLeuGlnIleSerGlnAspGlnAlaLeuLeuTyrAsnTyr 457  
 Db 1367 CACTGCATCTCACTAGAAATTTCAAGAGTAAGGCTCTTTGTATATATAT 1417

RESULT 2  
 AAQ38415  
 ID AAQ38415 standard; DNA; 1636 BP.  
 XX  
 AC AAQ38415;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 05-JUL-1993 (first entry)  
 XX  
 DE Polylacturonase cDNA clone pTOMC.  
 XX  
 KW polylacturonase; PG; pectin esterase; PE; expression regulation;  
 KW fruit softening enzymes; flowering plants; fruiting plants;  
 KW antisense RNA.  
 XX  
 OS Lycopersicon esculentum.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 51..142  
 FT /\*tag= a  
 FT /label= polylacturonase (PG)  
 XX  
 PN EF532060-A1.  
 XX  
 PD 17-MAR-1993.



XX 06-NOV-1987; 92EP-00117411.  
XX  
XX 11-NOV-1986; 86GB-00026879.  
XX  
XX (ICIL) IMPERIAL CHEM IND PLC.  
XX (ZENB) ZENECA LTD.  
XX  
XX Bldges IG, Grieron D, Schuch MW;  
XX  
XX WPI; 1993-087084/11.  
XX P-PSDB; AAR32107.  
XX  
XX Recombinant DNA for flowering and fruiting plants e.g. tomatoes ripening  
XX PT control - comprises base sequence for transcription contg. inverted  
XX PT sequence of bases complementary to bases in anti sense ribonucleic acid  
XX encoding softening enzymes, or gene expression regulation.  
XX  
XX Example 12; Fig 1; 20pp; English.

XX This is the sequence of clone pTOM6 which contains the coding sequence of  
XX CC polygalacturonase (PG). It was used to isolate the PG promoter in the  
XX CC construction of a vector encoding antisense RNA to the PG cDNA and PG  
XX CC gene. This would be useful to regulate the expression of the fruit  
XX CC softening enzymes in flowering and fruiting plants. Such antisense RNA  
XX CC would delay fruit softening. (Updated on 25-MAR-2003 to correct PN  
XX CC field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-  
XX CC 2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)  
XX

SO Sequence 1636 BP; 598 A; 232 C; 290 G; 516 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 6, 09e-219 Length: 1636  
Score: 2390.00 Matches: 457  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-691-374-2 (1-457) x AAO38415 (1-1636)

QY 1 MetValIleGlnArgAsnSerIleLeuLeuIleIlePheAlaSerSerIleSer 20  
DB 51 ATGGTATCCAAAGAAAGATGATTCCTTCATTAATTAATTTCTTCATCAATTTCA 110  
QY 21 ThrCysArgSerSerValIleAspAspAsnLeuPheLeuGlnValTyrAspAsnIleLeu 40  
DB 111 ACTGTGAGAAAGCAATGTTATGACAAATTTATTCACAAAGTTATGATATATCTT 170  
QY 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerIleAsnIleGlu 60  
DB 171 GAAACAAGAAATTCCTCATGATTTTCAAGCTTATCTTATTTGAGCAAAATATTTGA 230  
QY 61 SerAsnAsnAsnIleAspIleValAspIleAsnGlyIleLeuValIleAsnValIleAsn 80  
DB 231 AGCAACAATTAATTAACAAGGTGATTAATAAATGGAATTAATGATTAATGACTTAC 290  
QY 81 PheGlyAlaLeuSerIleAspGlyLeuTyrTyrAspAsnIleAlaPheGlnGlnAlaTyrAsn 100  
DB 291 TTTGAGCTTAAGGCTATGAGAAACATATGATTAATTCATTGAGCAAGCATGGAAT 350  
QY 101 GluAlaCysSerSerArgTyrProValGlnPheValValProIleAsnIleAsnIleLeu 120  
DB 351 GAAGCATGTTCACTGAAACACCTGTTCAATTTGTGTTCTTAAACAAAGAAATTAATCTT 410  
QY 121 LeuLeuGlnIleThrPheSerGlyProCysArgSerSerIleSerValIlePheGly 140  
DB 411 CTCAGCAAAATCACCTTTCCAGTCCATGCAAGATCTTCTAATTCAGTAAAGATTTTGA 470  
QY 141 SerLeuGlnAlaSerSerIleSerAspTyrIleAspArgIleLeuTyrIleAlaPhe 160  
DB 471 TCCTTGAAGCACTGTAATAATTTCAAGTCAAAAGATTAAGAGGCTTTGATTCCTTT 530

QY 161 AspSerValGlnAsnLeuValValGlyGlyGlyTyrIleAsnGlyAsnGlyGlnVal 180  
DB 531 GATAGTGTTCAAAATTTAGTTGTGAGAGAGGAACATATCAATGACATGACAAAGTA 590  
QY 181 TPTTPProSerSerCysLeuIleAsnIleSerLeuProCysArgAspAlaProThrAla 200  
DB 591 TGTGTGCCAAGTCTTCGCAAAATTAATAATCATCTGCCATGACGAGGATGCACCAACGGCC 650  
QY 201 LeuThrPheTyrAsnCysIleAsnLeuIleValAsnLeuIleAsnIleLeuIleAsn 220  
DB 651 TTAACCTTCTGAAATTTGCAAAATTTGAAGTAATATCTAAGAGTAATAAATGACACA 710  
QY 221 GlnIleHisIleLeuPheGluSerCysThrAsnValValAlaSerAsnLeuMetIleAsn 240  
DB 711 CAAATTCATTCAAATTTGAGTCATGCACTAAATGTTGAGCTTCAATTTGATGATCAAT 770  
QY 241 AlaSerAlaIleSerProAsnThrAspGlyValHisAlaSerAsnThrGlnTyrIleGln 260  
DB 771 GCTTCAGCAAGAGCCCAAAATCTGATGAGTCCATGTATCBAATATCTCAATATTTCA 830  
QY 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280  
DB 831 ATATCTGATACATTAATTTGAAACAGATGATGATTTGATTTCAATTTCTGATCTCA 890  
QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyIleGlyIleSerIleGlySerLeu 300  
DB 891 AATGTGCGAGGCCCAAAATTAATCTTGCTGTCAGGTCAGTAAAGTAATGTAAGCTTA 950  
QY 301 GlySerGlyAsnSerGlyAlaTyrValSerAsnValThrValAsnGlnAlaIleIle 320  
DB 951 GATCTGCAAAATTCAGAACTTATGCTATATGTTACTGTAATTAAGACCAAAATTAATC 1010  
QY 321 GlyAlaGluAsnGlyValArgIleLeuTyrTPGInGlyGlySerGlyGlnAlaSerAsn 340  
DB 1011 GGTGCCGAAAATGAGAGTTAGATCAAGACTTGCGAGGAGATCTGCAAGCTAGCAAC 1070  
QY 341 IleLeuPheLeuAsnValGluMetGlnAspValValTyrProIleIleIleAspGlnAsn 360  
DB 1071 ATCAAAATTTCTAAATGTGAATGCAAGACGTTAAGTATCCATTAATTAATGACCAAAAC 1130  
QY 361 TyrCysAspArgValGluProCysIleGlnIlePheSerAlaValGlnValIleAsnVal 380  
DB 1131 TATTTGATCGAGTTGAAACATGATTAACAAGCTTTTCAGCATTCAGTGAAGAAATGTG 1190  
QY 381 ValTyrGluAsnIleLeuGlyThrSerAlaThrLeuValAlaIleLeuPheAspCysSer 400  
DB 1191 GTGTATGAATAATATCAAGGCGACAAGTGCACAAAGGTGCGCATTAATTTGATTCACGC 1250  
QY 401 ThrAsnPheProCysGluGlyIleIleMetGluAsnIleAsnLeuValGlyGlySerGly 420  
DB 1251 ACAAACTTTCCATGTGAAGAAATTAATGAGAAATTAATTAATTTAGTAGGGGAAGTGA 1310  
QY 421 LysProSerGlyAlaThrCysIleAsnValHisPheAsnAsnAlaGluHisValThrPro 440  
DB 1311 AAAACATCAGAGGCTACGTCGCAAAATGTCATTTTAACAAGTCTGAACATGTTACACA 1370  
QY 441 HisCysThrSerLeuGlnIleSerGluAspGlyAlaLeuLeuTyrAsnTyr 457  
DB 1371 CACTGCACTTCACTAGAAATTTCAAGGATGAAGCTCTTTGTGATTAATTAAT 1421  
RESULT 3  
AAC84653  
ID AAC84653 standard; cDNA; 1636 BP.  
XX AAC84653;  
AC  
XX  
DT 20-APR-2001 (first entry)  
XX  
XX cDNA sequence of pTOM6 encoding a polygalacturonase (PG) enzyme.  
XX Pectin; pectin methyltransferase; PMT; polygalacturonase; PG; tomato;  
KM food product; yogurt; milk; fruit juice; whey drink; de-esterification;  
KM pTOM6; SS.

XX OS Lycopersicon esculentum.  
XX FH Key Location/Qualifiers  
XX FT CBS 51.1424  
XX FT /tag= a  
XX FT /product= "PG enzyme"  
XX PN MO200078982-A1.  
XX PD 28-DEC-2000.  
XX PF 15-JUN-2000; 2000MO-IB000869.  
XX PR 17-JUN-1999; 99GB-00014209.  
XX PA (DANI-) DANISCO AS.  
XX PI Christensen TME, Krelberg JD;  
XX DR WPI; 2001-091573/10.  
XX DR P-PSDB; AAB48338.  
XX PT Modifying pectin, for foodcuffs preparation, involves transforming host  
XX PT having pectin methyltransferase (PME) and polygalacturonase (PG) activity by  
XX PT silencing PG activity, to increase PME to PG ratio.  
XX PS Claim 6; Fig 1; 78pp; English.  
XX PS  
XX CC The invention provides a new method for modifying pectin that involves  
XX CC providing a host having pectin methyltransferase (PME) activity and  
XX CC polygalacturonase (PG) activity, transforming the host by silencing PG  
XX CC activity to provide an increased PME to PG ratio, preparing a PME extract  
XX CC from the transformed host, and using the PME extract to modify pectin. A  
XX CC PME modified pectin is useful for foodcuffs preparation, and to impart  
XX CC an increased functionality to food products such as yogurt, milk/fruit  
XX CC juice and whey drinks. PME is useful to reduce the number of ester groups  
XX CC in a pectin in a block-wise manner, and to de-esterify two or more  
XX CC adjacent galacturonic acid residue of a pectin on at least substantially  
XX CC all of the pectin chains. The present sequence represents the cDNA  
XX CC sequence of pTOM6 encoding a PG enzyme. A nucleotide sequence (seq id No.  
XX CC 3) encoding the PG enzyme which is deposited as pTOM23 with NCIMS  
XX CC (Accession number 12373) is also claimed. The sequence for seq id No. 3  
XX CC is not provided in the specification  
SQ Sequence 1636 BP; 598 A; 232 C; 290 G; 516 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 6.09e-219 Length: 1636  
Score: 2390.00 Matches: 457  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-10-691-374-2 (1-457) x AAC84653 (1-1636)  
QY 1 MetValIleGlnArgSerSerIleuLeuIleIleIlePheAlaSerSerIleSer 20  
DB 51 ARGSTATTCAGAGAAATGATATCTCCCTTCATTAATTAATTTGGCTTCATCAATTTCA 110  
QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheLysGlnValIleAspAsnIleLeu 40  
DB 111 ACTTGTAGAGCAAGTATTAATGATGCAATTAATTTCAACAAGTTTATGATTAATTCCTT 170  
QY 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerTyrAsnIleGlu 60  
DB 171 GAACAAGAAATTTGCATGATATTTTCACCTTATCTTTTATTTAGACAAAAAATATTGAA 230  
QY 61 SerAsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSer 80  
DB 231 AGCAACATAATTAATTTGACAGAGTTGATTAATAATGAGATTAAATGATTAATTAATTAAC 290

QY 81 PheGlnValAlaLysGlnLysArgLysThrTyrAspAsnIleAlaPheGlnGlnAlaTyrAsn 100  
DB 291 TTGGAGCTTAAGGCTAGTAAAGAAAAACATATGATTAATTAATGATTTTGGACCAACATGAAAT 350  
QY 101 GluAlaCysSerSerArgThrProValGlnPheValAlaProLysAsnLysAsnTyrLeu 120  
DB 351 GAAGCATGTTCACTCAAGAACCTGTTCAATTTGGTTCCTTAATAACAGAAATATATCTT 410  
QY 121 LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly 140  
DB 411 CTCAGCAAAATACCTTTTCAGCTTCATGCAAGATCTTCTAATTCAGTAAGATTTTGTGA 470  
QY 141 SerLeuGlnAlaSerSerLysIleSerAspTyrLysAspArgArgLeuTyrIleAlaPhe 160  
DB 471 TCCTTAGAAGCATCTAGTAATAATTTCAAGCTCAAGATGAAGAGCTTTGAGATGCTTTT 530  
QY 161 AspSerValGlnAsnLeuValAlaGlyGlyGlyThrIleAsnGlnLysGlnVal 180  
DB 531 GATAGTGTCAAAATTTAGTTGTTGAGAGAGAGAACTATCAATGCAATGCAAGCAAGTA 590  
QY 181 TrpTyrProSerSerCysValIleAsnLysSerLeuProCysArgAspAlaProThrAla 200  
DB 591 TGGTGGCAAGTTCTTGCAGAAATAAATTAATCACTGCGCATGACAGGATGACCAACGCGC 650  
QY 201 LeuThrPheTyrAsnCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsnAlaGln 220  
DB 651 TTAACTTCTGGAATTGCAAAAATTTGAAAGTGAATATCTAAAGCTAATAAATGACAA 710  
QY 221 GlnIleHisIleLysPheGlnLysSerCysThrAsnValAlaAlaSerAsnLeuMetIleAsn 240  
DB 711 CAATTCATATCAAAATTTGATGTCATGACCTAAATGTTGTACTTCAAAATTTGATGATCAAT 770  
QY 241 AlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260  
DB 771 GCTTCAGCAAAAGAGCCCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 830  
QY 261 IleSerAspThrIleIleGlyThrGlnAspAspCysIleSerIleValSerGlySerGln 280  
DB 831 ATATCTGATATCTAATTAATTTGAGACAGGATGATGATGATGATGATGATGATGATGATGAT 890  
QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300  
DB 891 AATGTGACAGCCCAAAATATTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 950  
QY 301 GlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaLysIleIle 320  
DB 951 GGAATCGGAATTCAGAACTTAATGTCTAATGTTAAGTAATGAAAGCCAAATTTATTC 1010  
QY 321 GlnAlaGlnAsnGlyValArgIleLysThrTyrGlnGlyLysSerGlyGlnAlaSerAsn 340  
DB 1011 GGTGGCGAAATAGACTTGGATGATCAAGACTTGGACGGAGGATGCGCAAGCTAGCAAC 1070  
QY 341 IleLysPheLeuAsnValGlnMetGlnAspValLysTyrProIleIleIleAspGlnAsn 360  
DB 1071 ATCAAAATTTCTGAATGCGAAATGCAAGCGTTAAGTATCCCATTAATTAAGCAACAAAC 1130  
QY 361 TyrCysAspArgValGlnProCysIleGlnGlnPheSerAlaValGlnValLysAsnVal 380  
DB 1131 TATTTGATCGAGTTGACATGATATACAAACAGTTTTCAGCAAGTTTCAAGTAAAGTAAGTG 1190  
QY 381 ValTyrGlnAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAspCysSer 400  
DB 1191 GTGTATGAGAAATATCAAGGGCAAGTGCCAAAGAGTGGCCATTAATTTGATTTGACGC 1250  
QY 401 ThrAsnPheProCysGlnGlyIleIleMetGlnAsnIleAsnLeuValGlyGlnSerGly 420  
DB 1251 ACAAACCTTTCATGTGAAGGAATTAATATGAGAAATTAATTAATTAATTAATTAATTAATTA 1310  
QY 421 LysProSerGlnAlaThrCysLysAsnValHisPheAsnAsnIleGlnHisValThrPro 440  
DB 1311 AAACCATCAAGAGCTACGTGCAAAAATGCTCATTTTAAACATGCTGAACATGTTTAAACCA 1370  
QY 441 HisCysThrSerLeuGlnIleSerGlnAspGlnAlaLeuLeuTyrAsnTyr 457

```
Db      1371 CACTGCACTTCAAGAAATTCAGAGATGAAGCTCTTTGTATATTTAT 1421
RESULT 4
AAN80487
ID      AAN80487 standard; DNA; 1637 BP.
XX
XX      AAN80487;
XX
XX      25-MAR-2003 (revised)
DT      20-NOV-1990 (first entry)
XX
XX      Plasmid clone pTom6.
DE
XX      Fruit ripening; polygalacturonase; pectin esterase; ss.
XX
XX      Synthetic.
OS
XX      Key      Location/Qualifiers
FH      CDS      51..1424
FT      /tag= a
FT      /label= polygalacturonase
XX
XX      EP271988-A.
XX
XX      22-JUN-1988.
XX
XX      06-NOV-1987; 87EP-00309853.
XX
XX      11-NOV-1986; 86GB-00026879.
XX
XX      (ICIL ) IMPERIAL CHEM IND PLC.
XX      (ZENB ) ZENBICA LTD.
XX
XX      Bridges IG, Schuch WW, Grierson D;
XX      WPI: 1988-169271/25.
XX      P-PSDB; AAP80299.
XX
XX      Recombinant DNA comprising promoter and terminator sequences - useful in
XX      plants for altering ripening properties esp. in tomatoes.
XX
XX      Disclosure; Page ?; 22pp; English.
XX
XX      This cDNA clone, encoding polygalacturonase (PG), is used to produce
XX      antisense mRNA (with an inverted sequence to that of PG mRNA) which is
XX      inserted into a vector used to transform plants which thereafter have
XX      altered ripening properties. The inverted sequence and the PG mRNA form a
XX      double-stranded structure which inhibits expression of the PG mRNA. See
XX      also AAN80488. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX      Sequence 1637 BP; 598 A; 232 C; 290 G; 517 T; 0 U; 0 Other;
SQ
SQ
Alignment Scores:
pred. No.: 6.09e-219 Length: 1637
Score: 2390.00 Matches: 457
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-10-691-374-2 (1-457) x AAN80487 (1-1637)
QY      1 MetValIIGlnArgAsnSerIleLeuLeuIleIleIlePheAlaSerSerIleSer 20
Db      51 ATGGTATATCCAAAGGATAGTATTCCTCTCATATTATTTTGGTTCATCAATTTTCA 110
QY      21 ThCysArgSerAsnValIleAspAspAsnLeuPheLeuGlnValIleAspAsnIleLeu 40
Db      111 ACTTGAGAGAGCAATGTTATGTACAAATTTATTCAAACAAAGTTTATGATATATCTT 170
QY      41 GlnGlnGlnPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerIleAsnIleGln 60
```

```
Db      171 GAACAAGAAATTTGCTCATGATTTTCAAGCTTATCTTCTTATTTGACAAAAATATGAA 230
QY      61 SerAsnAsnAsnIleAspIleValIlePheAsnGlnIleLeuValIleAsnValIleuSer 80
Db      231 AGCAACAATTAATATATGACAAAGGTGTATAAAATGGATTTAAAGTATTAATCTACTTACG 290
QY      81 PheGlnAlaLeuGlyAspGlyIleThrIleAspAsnIleAlaPheGlnGlnIleIlePheAsn 100
Db      291 TTTGAGCTTAAGGGTGTATGAAAAACATATGTATTAATATGCAATTTGAGCAAGCATGGAAT 350
QY      101 GlnAlaCysSerSerArgThrProValGlnPheValIleProLeuAsnIleAsnIleuLeu 120
Db      351 GAGCATGTTCATCTAGAAACACCTGTCAATTTGGTTCCTTAAACAAGAAATTAATCTT 410
QY      121 LeuIleGlnIleThrPheSerGlyProCysArgSerSerIleSerValIlePheGly 140
Db      411 CTCAGCAAAATCACCTTTTCAGGTCCAGATGCATGATCTTCTATTTCAATTAAGATTTTGGCA 470
QY      141 SerLeuGlnAlaSerSerIleIleSerAspTyrIleAspAspArgLeuTrpIleAlaPhe 160
Db      471 TCCTTAGAAGCATCTAGTAAATTTTCAGCTCAAAAGATAGAAAGGCTTTGGATGCTTTT 530
QY      161 AspSerValGlnAsnLeuValIleGlyGlyGlyThrIleAsnGlyAsnGlyGlnVal 180
Db      531 GATAGTGTTCAAATTTTACTTGTGGAGAGAGAGACATCATATGCAATGCAATGCAACAGTA 590
QY      181 TrpTrpProSerSerCysLeuIleAsnIleSerLeuProCysArgAspAlaProThrAla 200
Db      591 TGTGGCCAAAGTTCTTGCAAATAATTAATATCATGTCATGCAAGGATGCAACACGCGCC 650
QY      201 LeuThrPheTrpAsnCysLeuAsnLeuIleValIleAsnLeuIleSerIleAsnIleGln 220
Db      651 TTAACCTCTTGAAATTTGCAAAATTTGAAAGGAAATTAATCTAAGAGTAAATTAATGACAA 710
QY      221 GlnIleHisIleLeuPheGlnIleSerCysThrAsnValIleAlaSerAsnLeuIleAsn 240
Db      711 CAATTCATATCAATTTTATGATCATGCACTAATGTTGTAGCTTCAAAATTTGATGATCAAT 770
QY      241 AlaSerAlaLeuSerProAsnThrAspGlyValIleValIleSerAsnThrGlnTyrIleGln 260
Db      771 GCTTCGCAAAAGGCCAAATATCTGATGAGTCCATGTATCAAAATCTCAATATATTTCA 830
QY      261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValIleSerGlySerGln 280
Db      831 ATATCTGATATCAATTTATGAAACAGGTATGATGATGATTCATATGTTCTTGATCTCA 890
QY      281 AsnValGlnAlaThrAsnIleThrCysGlyProGlnIleGlyIleSerIleGlySerLeu 300
Db      891 AATGTCCAGGCCAAATATTTTCTTGTGTCAGAGTCATGTAATAGTATTTGGAAGCTTA 950
QY      301 GlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaLeuIleIle 320
Db      951 GGATCTGCAAAATTCGAAGACTTATGTGTATGTTACTGTAAATCAACCAAAATTTATTC 1010
QY      321 GlyAlaGlnAsnGlyValArgIleLeuSerThrTrpGlnGlyIleSerGlyGlnIleAspAsn 340
Db      1011 GGATGCCGAAAAGAGGTTATGATGATCAAGACTTGTGCGAGGAGATCTGGAACAGCTACCAAC 1070
QY      341 IleLeuPheLeuAsnValGlnIleCysGlnIleAspValIleTyrProIleIleIleAspGlnAsn 360
Db      1071 ATCAAAATTTCTGAATGTGGAATTCGAAGACGTTAAGTATCCATATTTATAGCCAAAC 1130
QY      361 TyrCysAspArgValGlnIleProCysIleGlnIlePheSerAlaValGlnValIleAsnVal 380
Db      1131 TATTTGATCGAGTTGAACCATGTATACAAAGTTTTCGAGCTTCAAGTGAATAATGTG 1190
QY      381 ValIleTyrGlnAsnIleLeuGlyThrSerAlaThrIleValAlaIleLeuPheAspCysSer 400
Db      1191 GTGTATGAATAATATCAAGGCGCAAGTGCACAAAGGTGGCCATTAATATTTGATTCACAC 1250
QY      401 ThrAsnPheProCysGlnGlyIleIleIleIleCysAsnIleAsnLeuValGlyIleuSerGly 420
Db      1251 ACAAACTTTCCATGTGGAAGATTTAATGAGAAATATTAATTTTGTAGGGGAAAGTGA 1310
```



Db	2227	AAGCGGCGCATTAATAATTGGATTGACGACACAACTTCCATGTGAAAGAAATTATATATGAG	2285		
Qy	412	AsnIleAsnLeuValGlyGluSerGlyIysProSerGluAlaThrCysAlaSerValHis	431		
Db	2287	AATATTAATTATTAAGTAGGGGAAAGTGGAAACCATCAGAGGCTACGTGCATAAATGTCAT	2346		
Qy	432	PheAsnAsnAlaGluHisValThrProHisCysThrSerLeuGluHisSerGluAspGlu	451		
Db	2347	TTTAAACAATGCGCAACATGTTACACCACTGCACCTTCACTAAGAAATTTCCAGAGATGAA	2406		
Qy	452	AlaIleuLeuTyrAsnTyr	457		
Db	2407	GCTCTTTGTATTAATTAT	2424		
RESULT 6					
ID	ADG87885				
AC	ADG87885	standard; cDNA; 1296 BP.			
XX	ADG87885;				
XX	22-APR-2004	(first entry)			
DE	A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #327.				
XX	Pathogen infection-related gene; plant; Peronospora parasitica;				
KM	defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant;				
KW	oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.				
XX	Arabidopsis thaliana.				
OS					
XX	MO200222675-A2.				
PN	21-MAR-2002.				
PD	14-SEP-2001; 2001MO-US028506.				
PE	15-SEP-2000; 2000US-0232778P.				
PR	22-JUN-2001; 2001US-0300183P.				
XX	(SYGN ) SYNGENTA PARTICIPATIONS AG.				
PA	(LYNC-) UNIV NORTH CAROLINA.				
PA	(GLAZ/) GLAZEBROOK J.				
PA	(WANG/) WANG X.				
PA	(DANG/) DANG J L.				
PA	(EULG/) EULGEM T.				
PA	(ZHUT/) ZHU T.				
XX	Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;				
PI	WPI; 2002-292409/33.				
DR					
PT	Novel isolated polynucleotide, useful for conveying pathogen resistance				
PT	to plants, and for identifying plants infected with a pathogen.				
XX					
PS	Claim 3; SEQ ID NO 327; 605bp; English.				
XX					
CC	The invention relates to 691 Arabidopsis thaliana genes (ADG87559--				
CC	ADG87557)) whose expression is altered in response to pathogen infection,				
CC	and to homologues of these genes from other plants or fungi, especially				
CC	from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),				
CC	cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The				
CC	expression of genes of the invention was upregulated or downregulated in				
CC	Arabidopsis plants infected with the oomycete Peronospora parasitica,				
CC	indicating that they play a role in defence mechanisms. The genes of the				
CC	invention are regulated by RPP7 or RPP8 which act via unconventional				
CC	signalling cascades, or by the RPP4-dependent pathway. The invention also				
CC	relates to polypeptides encoded by the pathogen infection-related genes;				
CC	promoter motifs from pathogen infection-related genes (ADG888243-ADG88327)				
CC	; expression cassettes, host cells and pathogen-resistant transgenic				
CC	plants and their progeny comprising a polynucleotide of the invention;				
CC	and a method of identifying a plant cell infected with a pathogen. The				
CC	polynucleotide sequences and methods of the invention are useful for				
CC	identifying plants infected with a pathogen, and for conferring				

CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses, nematodes and insects (e.g., aphids). The present sequence represents an Arabidopsis thaliana gene whose expression is altered in response to CC peronospora parasitica infection. Note: The sequence data for this patent CC did not form part of the printed specification, but was obtained in CC electronic format directly from WPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1296 BP; 424 A; 244 C; 293 G; 335 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,98e-98	Length:	1296
Score:	1127.00	Matches:	223
Percent Similarity:	65.33%	Conservative:	71
Best Local Similarity:	49.56%	Mismatches:	116
Query Match:	47.15%	Indels:	40
DB:	6	Gaps:	6

US-10-691-374-2 (1-457) x ADG87685 (1-1296)

```
QY      7 SerIleuLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCyAArgSerAenVal 26
      25 GCGTGTTCCTATGAGCGTCTCTTGATGCTGCTGTTGTAAAGCTTTGATGCAACGTT 84
QY      27 IleAap-----AaspAenuePheArgVal 35
      85 GATGATGGATATGTCATGAGATGAGAGACTTTGAATCCGATAGCTTGCTCAAGCTCAAC 144
QY      36 TyrAaPAnIleLeuGlnGlnIupPheAlaHlaaPheGlnAlaTyrLeuSerTyrLeu 55
      145 AACGATGATGTTCTT-----AGCTTGATTAACCTCGAC 177
QY      56 SerIyAaenIleGlnSerAaPAnAnIleAapLyVaIaPlyAaenGlyIleVal 75
      178 GAACACCATTTGGAAAGCATCA-----ACC 201
QY      76 IleAaenValLeuSerPheGlyAlaIyArgIyAaPlyThrTyrAaPAnIleAlaPhe 95
      202 GTTAACTGTTCAACCTTCGAGCCAAAGAGATGAGAAACATGATATACTCAAGCCCTTC 261
QY      96 GluGlnAlaTyrAaenGlnIaCySerSerAArgThrProValGlnPheValProlys 115
      262 AAGAAAGCATGAGAAAGAAAGCATTTCAACAATGAGATTACTACTTCTTGTTCCATAA 322
QY      116 AaenLyAaenTyrLeuLysGlnIleThrPheSerGlyProCyAArgSerSerIleSer 135
      322 GGAAGAGACTTATCTCTTAAAGTCTACTCGATTAGAGGCCATGCAATCCTTAACGTAAAC 381
QY      136 ValLyIlePheGlySerLeuGlnAlaSerSerLyIleSerAapLyTyrAaPAnArg 155
      382 TTTCGATCTCTAGGACCTTATCAGCATCTACGAAACGTTCAAGATTACAAAGCAAAAC 441
QY      156 LeuTPRIleAlaPheAaPserValGlnAaenLeuValIaGlyGly-----GlyThr 173
      442 CATTCGCTTATCTTGAAGAGACGTTAAACAATCTATCAATGACAGCGTGCAGCGGAATT 501
QY      174 IleAaenGlyAaenGlyGlnValTyrTPRProSerSerCyAlysIleAaenLySerLeuPro 193
      502 ATTATATGCAAGCAAGAAACCTGCTGCGAAGACTCATGCAAAATTCGCAAAATTCAGCCA 561
QY      194 CyAaTgaPAnIleProThrAlaLeuThrPheTPRAnCyAlyAaenLeuLyValAaenAn 213
      562 TGCACAAAAAGCTCCAAACGGCTCTTACTTATCAATTTAAAGAAATTTGAATGCAAGAAAT 621
QY      214 LeuLySerTyrAaenAlaGlnGlnIleHlaIleLyPheGlnSerCyThrAaenVal 233
      622 CTGAGGGTGCAAAATGCGCAGCATTCAGATTTCATTGACAAATGCAACAAAGTTGAA 681
QY      234 AlaSerAaenIleAaenIleAaenAlaSerAlaLySerProAaenThrAaPlyValHlaVal 253
      682 GTTATGTAATGTTGAAGATCACTGCTCGGGCGATAGTCCCAACACAGATGTATTCATATC 741
QY      254 SerAaenThrGlnTyrIleGlnIleSerAapThrIleIleGlyThrGlyAaPAnCySile 273
```



```
QY 234 AAlaSerAsnLeuMetIleAsnAlaSerAlaIleSerProAsnThrAspGlyValHisVal 253
Db 682 GTTAGTAATGTTAGATCACTGCTCCGGGCGATAGTCCCAACAGATGATATCCATATC 741
QY 254 SerAsnThrGlnTyrlleGlnIleSerAspThrIleGlyThrGlyAspAspCysIle 273
Db 742 ACTAATATCTCAAAACATTGAGCTCCACATCTCAATATGAGAACAGTGTATTTTATA 801
QY 274 SerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHis 293
Db 802 TCCATTGAGAGATCGAAGCGCAAAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCA 861
QY 294 GlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaThrValSerAsnValThr 313
Db 862 GGGATACGATTGGAGCTTGGGAGCGACAAATTCAGAACCTTATGCTCGGGATTAAT 921
QY 314 ValAsnGlnAlaIleValIleIleGlyValGlnValArgIleIleThrGlnGlnGly 333
Db 922 GTGGATGAGTGTAAAGTTCTCTGAGATGACAAATGAGATTAGATTAAAGCTTATCAGGA 981
QY 334 GlySerGlyGlnAlaSerAsnIleIleIleValAsnValGlnMetGlnAspValIleTy 353
Db 982 GGATTCAGAACTGCCAAGAACATTAAATTTCAAAATTTCAATTTCAATTTCAATTTCA 1041
QY 354 ProIleIleIleAspGlnAsnTyIleCysAspArgValGluProCysIleGlnIlePheSer 373
Db 1042 CCGATATATATCAGCAAGGACTATCTGCGACAG--GACAAATGCGAAGAACCAAGAGTGG 1098
QY 374 AAlaValGlnValIleAsnValValTyIleGlnAsnIleIleGlyThrSerAlaThrIleVal 1098
Db 1099 GCACTGCAAGTGAAGAACCTTGTGTACAGACATATCTGTGTGAGAGCTGAGAGTGTG 1158
QY 394 AlaIleIleValPheAspCysSerThrAsnPheProCysGlnGlyIleIleMetGlnAsnIle 413
Db 1159 GCGATTACCTTGAATTGACGAGAGATATCTCATGTCAGAGGATTTGTGCTTGAGAACGTG 1218
QY 414 AsnLeuValGlyIleSerGlyIleProSerGlnAlaThrCysIleValAsnValHisPheAsn 433
Db 1219 AAAATTAAGGA-----GGAACAGCTTCTTGCAAAATGCGCAATGTTAA 1263
QY 434 AsnAlaGlnHisValIleThrProHisCysThr 443
Db 1264 AATCAAGGACCGCTTCTCTCAATATGCTCT 1293
RESULT 8
AAC46786 standard; DNA; 1393 BP.
ID AAC46786
XX
AC AAC46786;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 51398.
XX
KW Hybridization assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132483P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
```







```

Db 397 ACCTATCCGTCATCAAAAACGATCGATACAAAGATATCAGCAATGATTAATGTTT
Qy 161 ApsSeraValGlnAsnLeuValGlyGlyGly-----GlyThrIleAsnGlyAsnGly 178
Db 457 GATGGCGTTAAACATCTATCATCGATGAGGCGACATCGGGGTTGTGAGCGAAGACCGC 516
Qy 179 GlnValTrpTrpProSerCysIleAsnLeuSerLeuProCysAraAphAlaPro 198
Db 517 GAACCTGGTGGCAAACTCATGCAACCGAACAAGCT-----555
Qy 199 ThrAlaLeuThrPheTrpAsnCybIysAsnLeuValAsnAsnLeuSerIysAsn 218
Db 556 AAGGCTCTTACTTTTACCACTCGAAAGCTGATGATGAGAACTGAAGCTGAGAAAT 615
Qy 219 AlaGlnGlnIleHisIleLeuPheGluSerCysThrAsnValAlaSerAsnLeuMet 238
Db 616 GCACAGCATCATCAATTCATGTTCAATTAATAATGCTCAACGTCAGTCTTAATGCTGTG 675
Qy 239 IleAsnAlaSerAlaIysSerProAsnThrAraPglValHisValSerAsnThrGlnTyr 258
Db 676 GTAACTGGCGCTCGGATAGTCTTAACACCATGTTATCATATCATCAACCCCAAAAC 735
Qy 259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly 278
Db 736 ATTGAGTCTCCGAATTCATATGGAACAGCGCATGATGATCATATCATTAAGAAAGTGA 795
Qy 279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyValIleSerIleGly 298
Db 796 TCACAAATGTTCAAAATGATGATATTAATCTGCGGTCGCGGTCAAGATCAGATATGGG 855
Qy 299 SerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaIys 318
Db 856 AGCTTGGAGATGACAAATTCAAAGCTTTGTCTCAGGCGTGACTGTGATGCTGCTAAG 915
Qy 319 IleIleGlyAlaIleAsnGlyValArgIleLeuThrTrpGlnGlySerGlyGlnAla 338
Db 916 CTTTCGGTACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
Qy 339 SerAsnIleIysPheLeuAsnValGluMetGlnAspValIleTyrProIleIleLeuP 358
Db 976 AGCAATATTTATTTCAAAACATTCAGATGATTAATGATTAAGATTCGATCATTAATCGAC 1035
Qy 359 GlnAsnTyrCysAspAspValGluProCysIleGlnGlnPheSerAlaValGlnValIys 378
Db 1036 CAAAGCTACTGCAAG--AGCAAAATGCATCAAGAGAAATCCCGGTCCAAAGTGAAG 1092
Qy 379 AsnValValTyrGluAsnIleIysGlyThrSerAlaThrIleValAlaIleIysPheAsp 398
Db 1093 AACGTGTGTACCGGACATAGTGGCATAGCGCATCGAAACCAATTAACGTTTAAAC 1152
Qy 399 CysSerThrAsnPheProCysGluGlyIleIleMetGluAsnIleAsnLeuValGlyGlu 418
Db 1153 TGGAGCAAGAACTATCATGCAAGAAATTTGCTTCAAGAGTGAACCTTAAGAAA--- 1209
Qy 419 SerGlyIysProSerGlnAlaThrCysIleAsnValHisPheAsnAsnAlaIleHisVal 438
Db 1210 -----GGAAGCACTTCCACCAATGCTTAATGTGTTGATTAAGAGACGCTTT 1257
Qy 439 ThrProHisCysThrSer 444
Db 1258 CTGCTCAGTGCMACTCC 1275

```

RESULT 10  
AAT63603 standard; cDNA to mRNA; 1631 BP.  
XX AAT63603;

DT 17-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 08-JUL-1997 (first entry)  
XX  
DE Dehiscence zone-selective endopolygalacturonase cDNA.

```

XX Dehiscence zone; endo-polygalacturonase; seed shattering; oilseed rape;
KW transgenic plant; antisense; ribozyme; ds.
XX
OS Brassica napus; cv. Topaz.
FH
FH Key Location/Qualifiers
FT CDS
FT /tag= a
FT 95..136
FT sig_peptide
FT /tag= b
FT 164..1393
FT mat_peptide
FT /tag= c
FT primer_bind
FT /tag= d
FT /note= "primer PG1"
FT complement(884..900)
FT primer_bind
FT /tag= e
FT /note= "primer PG3"
FT 1059..1073
FT primer_bind
FT /tag= f
FT /note= "primer PG2"
FT 1229..1245
FT primer_bind
FT /tag= g
FT /note= "primer PG5"
FT /tag= h
FT /note= "base 1439 is given as n in the specification"
PN MO9713865-A1.
PD 17-APR-1997.
PF 04-OCT-1996; 96WO-EP004313.
PR 06-OCT-1995; 95EP-00402241.
PR 08-DEC-1995; 95EP-00203328.
XX
XX (PLBZ ) PLANT GENETIC SYSTEMS NV.
PI Ulyskov P, Child R, Van Onckelen H, Prinsen E, Borchardt B;
PI Sander L, Petersen M, Bundgaard Poulsen G, Botterman J;
DR WPI, 1997-235901/21.
XX
XX Transgenic plant containing dehiscence zone selective chimeric gene - has
XX modified dehiscence properties, especially delayed pod dehiscence.
XX Claim 6; Page 48; 65pp; English.
XX
XX cDNA clone X (AAT63603) includes a coding sequence for oilseed rape
XX dehiscence zone (DZ)-selective endopolygalacturonase (PG). DZ cDNA was
XX subjected to PCR amplification using primers (see also AAT63605-08) based
XX on conserved regions of PG amino acid sequences. PG-related clones (see
XX also AAT63609-14) were identified, of which only PG35-8 was specific to
XX the DZ. This clone was used to screen a DZ-selective cDNA library prepd.
XX from poly-A+ RNA isolated from the DZ 6 wk after anthesis, yielding clone
XX X. Ipg35-8 was also used to screen a genomic library to identify the DZ-
XX selective PG gene promoter region (AAT63604). These sequences can be
XX utilised in novel chimeric genes to modify the dehiscence properties of
XX transgenic plants, partic. the pod dehiscence properties of B. napus.
XX This can improve the seed yield from the transformed plants. (Updated on
XX 25-MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to standardise
XX OS field)
XX
XX Sequence 1631 BP; 554 A; 291 C; 339 G; 446 T; 0 U; 1 Other;

```

## Alignment Scores:

Pred. No.: 1.5e-96 Length: 1631  
Score: 1113.00 Matches: 223  
Percent Similarity: 65.32% Conservative: 69  
Best Local Similarity: 49.89% Mismatches: 121  
Query Match: 46.57% Indels: 34



XX	Sequence	1657 BP; 541 A; 314 C; 351 G; 451 T; 0 U; 0 Other;
Alignment Scores:		
Pred. No.:	1,39e-95	Length: 1657
Score:	1103.00	Matches: 721
Percent Similarity:	65.54%	Conservative: 20
Best Local Similarity:	49.77%	Mismatches: 125
Query Match:	46.15%	Indels: 7
DB:	2	Gaps: 7
US-10-691-374-2 (1-457) x AATG33994 (1-1657)		
QY	7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThcCyArGSerAsnVal	26
DB	169 GCTATTTCCTTATGTGGTCTTTTGTANTGCTCGCTTCGCCAAGCTTGAAGTGAACGTA	221
QY	27 IleAsp-----AspAsnLeuPheLysGlnValTYrAspAsnIleGlnu	41
DB	229 GATGTGGATATGCTCATGAAGATGCAAGCTTCGAACC-----GATAGTTATATCAAG	285
QY	42 GlnGluPheAlaHisAspPheGlnAlaTYrLeuSerTYrLeuSerLysAsnIleGluSer	61
DB	283 CTCACCAACGACGACGACGCTTCACTTGAAGAGCTCCGATACCCATCAACCGAATCA	342
QY	62 AsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSerPhe	81
DB	343 TCA-----ACTGTTAGTTCGATTC	366
QY	82 GlnAlaLysGlyAspGlyLysThrTYrAspAsnIleAlaPheGlnGlnAlaTPrAsnGlnu	101
DB	367 GGAGCAAAAGGTATGAGAAAAACCGATGATCTCAGGCTTCAAGAAAGATGAGAGAC	426
QY	102 AlCysSerSerAsnThrProValGlnPheValAlaProLysAsnLysAsnTYrLeuLeu	121
DB	427 GCATGTTCAACAATGAGAGTGACTACTTTGTAATCTTAAAGGAAAGACTATATCTCTT	486
QY	122 LysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGlySer	141
DB	487 AAGTCTATATGATTCAGAGGCCCATGGAATCATTAAGTACGCTTCCAGATCCTAGGCACCT	546
QY	142 LeuGlnAlaSerSerLysIleSerAspTYr--LysAspArgLysLeuTPrIleAlaPhe	160
DB	547 TTATCAAGCTTCAAAAACGATCGATTCACATTAAGACAAACACACATCGGCTTATTTG	606
QY	161 AspSerValGlnAsnLeuValGlyGlyGly-----GlyThrIleAsnGlyAsnGly	178
DB	607 GAGGACGTTAATATATCATCATCATGATGCGGCTCGCGGAGATGTTGATATGCGAACGA	666
QY	179 GlnValTTrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaPro	198
DB	667 AAAATGTGGTGGCAAAATCATGCAAAATCGCAAAATCATTAACCATGCGCAAAAAGCGCA	726
QY	199 ThrAlaLeuThrPheTrpAsnCyLysAsnLeuLysValAsnAsnLeuLysSerLysAsn	218
DB	727 ACGGCTCTTACTCTGCACAACTTAACAATTTGAATGTGAAGAACTGAGACTGAAGAAAT	786
QY	219 AlaGlnGlnIleHisIleLysPheGluSerCysThrAsnValAlaLaserAsnLeuMet	238
DB	787 GCACGCAAGATTCAGATTCGATTGAGAAATGCAACAGTGTGATGATTAAAGATTTAAG	846
QY	239 IleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTYr	258
DB	847 ATCACTGCTCCGGGAGATATGTCACACACGATGATATTCATATTCGTGTCAATAAAC	906
QY	259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly	278
DB	907 ATTGAATCTCCAAATTCAGACATTTGGGCAAGATGATGATTCATATCATTCATGAAGATGA	966
QY	279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGly	298
DB	967 TCGCAAAATGTTCAAAATCAATGATTTAATCTTGGCGGCCCCGCTATGCGATCAAGATTTGA	1026

Oy	239	SerLeuGlySerGlyValSerGluAlaTyrValSerAsnValThrValAsnGluAlaLys	318
Db	1027	AGCTTGGGGGATGACAAATTCGAAAGCTTATGATCGGAATTAATGTGGATGCTACG	108
Oy	319	IleIleGlyAlaGluAsnGlyValArgIleLysThrTyrGlnGlyGlySerGlyGlnAla	338
Db	1087	CTTCTGAGACTGACCAATGAGATAGAAATCAAGACTTACAGGAGGCTCAGAACTGCT	114
Oy	339	SerAsnIleLysPheLeuAsnValGluMetGlnAspValLysTyrProIleIleLeuAsp	358
Db	1147	AAGAACAATTAAATTCGAAAGCAATTCGTATGAGATTAATGTCAGAAATCCGATCAATCGAC	120
Oy	359	GlnAsnTyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValLys	378
Db	1207	CAGAACTACTGACCAAG--GACAAATCGAACAACAAGATCTGCGGTTCAAGTGAAC	126
Oy	379	AsnValValTyrGluAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAsp	398
Db	1264	AAATGCTGATCGGAACATACAAAGGTACAGCGCAACGAGTGGCGATATGTTTAAAT	1322
Oy	399	CysSerThrAsnPheProCysGluGlyIleIleMetGluAsnIleAsnLeuValGlyLys	418
Db	1324	TGCAGTGTGAAATATCCATGACCAAGGTATTTGCTTGAAATGTCAACATTCGAAAGCA--	1380
Oy	419	SerGlyLysPheSerGluAlaThrCysLysAsnValHisPheAsnAlaGluHisVal	438
Db	1381	-----GGAAGACCTTCTTGCAAAATATCATGTGTTAAGATTAAGGCAACGTT	1428
Oy	439	ThrProIleCys	442
Db	1429	TCCTCTAATGTC	1440
RESULT 12			
AAZ22980	ID	AAZ22980	standard; DNA; 1657 BP.
AC	XX	AAZ22980;	
DT	XX	10-JAN-2000	(first entry)
DE	XX	Brassica napus Sac66 protein encoding DNA.	
KW	XX	Signal transduction protein; dehiscence; male sterile plant; DZ2 gene;	
KW	XX	shatter resistance; oilseed rape; Sac66 protein; ss.	
OS	XX	Brassica napus.	
PN	XX	W0949046-A1.	
PD	XX	30-SEP-1999.	
PF	XX	22-MAR-1999;	99WO-GB000905.
PR	XX	20-MAR-1998;	98GB-00006113.
PA	XX	(BIOG-) BIOGEMMA UK LTD.	
PI	XX	Wyatt P, Roberts JA, Whitelaw C;	
DR	XX	WPI: 1999-580449/49.	
DR	XX	P-PSDB; AAY42649.	
FT	XX	A nucleic acid encoding a signal transduction protein involved in plant dehiscence, useful for producing shatter resistant male sterile plants.	
PS	XX	Example 6; Fig 15; 71pp; English.	



-----COUNCIL OF CHURCHES 933

3 XX

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5.75e-87	1011.00	69.52%	53.28%	42.30%	1182	187	57	105	2	1

US-10-691-374-2 (1-457) x ADA70540 (1-1182)

```

QY 72 GYILEYVAIIIEAENVALIEUSERPHEGIALALYAGLYAASPGLYRTHRYAEP 91
DB 79 GGGAGTAAGTGTCTTACAGCATACAGCTACGGGGCTCAGAGAGGAGCATATATGAC 138
QY 92 AENILEAAPHGGLGALATTPANGIUAICYSESERARGTHRPROVALGHPHE 111
DB 139 ACCAAAGCATGGAGATACATGGGCTGCAGCTGCTCTCCAGCAAAACCTGCAGTTTG 198
QY 112 VALIAPROLYAENLYSANTYRLEUWLYSGINILETHRPHSEGLYPROCYARG 131
DB 199 CTATCTCCAGAGGAGCAAGAAATCTGATCAAGACACAAACTGCTGCTCCATGCAAA 258
QY 132 SETSERLESEVALYSEIIEPHEGLYSEIIEGLUALASERSELYSEIIESEAPTYR 151
DB 259 TCAAGCATCTCATGTGATGGTGAAGGCTATTGCTGCTTCCAGAGAGGTCAAGCTGG 318
QY 152 LYAAPARG-----ARGLEUTRPIIEAAPHSEPSERVALGIAENLEUVALIGLY 169
DB 319 ACCAAGAGACCATAAAGCAGCTGATTCGATCAGTGAGTGCTGCTGTTACTGCTACT 378
QY 170 GYLYGYLYTHRIIEANLYANGLYGINVALITRTPROSESERLYSEIIEAN 189
DB 379 GGTGTGGAGCATATAGTGAATGCAAGATTGGTGCAAAATTCATGCAAAACCAAC 438
QY 190 LYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLY 209
DB 439 TCAAGCTTCCATGCAAGAGCTCAAGGCTCAAGGCTCAAGCTTCAAGCTTCAAGGATCG 498
QY 210 LYFVALAENLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLY 229
DB 499 AAGGTGAGTATCTGAAAGTGTGAAACCCAGCAAAATTCATGCTGAGGAGTGC 558
QY 230 THRANVALALASERANLEUWETIEANALASERANLYSEIIEPHEGLYSEIIEPHEGLY 249
DB 559 ACCGATGTATGTGCTGCTGCTGATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 618
QY 250 GYVALIIEAVALSERANTHRGINTYRIIEGLIIESEAPRTHRIETIEGLYTHRGY 269
DB 619 GGAATTCATATCAAGCAGCAGCAGAGTGTGAGTACAGCTGATCAAGAGCCGG 678
QY 270 AAPAPCYEIISEIIEVALSERGLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLY 289
DB 679 GATGACTGATGTCATCAAGCAGCAGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 738
QY 290 GYLYPGGLYHAGLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLY 309
DB 739 GGAACCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 798
QY 310 SETSERVALIIEAVALSERANTHRGINTYRIIEGLIIESEAPRTHRIETIEGLYTHRGY 329
DB 799 AACATGTCAAGCTGCAAGATGTCAGTGTATGAGCAGCAGCAGCAGCAGCAGCAGCAG 858
QY 330 THRTGGLYLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLY 349
DB 859 ACATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 918
QY 350 AAPVALYETRYPROIIEIIEAEPGIAENTYCYAASAPARGVALGIUPCYSEIIE 369
DB 919 AATGTGTGAACCCATCATCATCATCATCAAAACTATGAGCTCTTCAACCCCTCAAG 978
QY 370 GINGINPHESEVALAVALGINVALYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLY 389
DB 979 CAACGAATCTGCAAGTGTGAGCAGTGTGCTCTTCAAGAACATCAAGGAGGAGGAGGAG 1038

```

QY 390 ALATHRYVAIIIEAENLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLY 409

DB 1039 GCATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1098

QY 410 METGIUANILEAENLYSEIIEPHEGLYSEIIEPHEGLYSEIIEPHEGLY 420

DB 1099 TTGGAGAAATGTCTCATCTCATCTCAAGGAGAGA 1131

RESULT 15

ADA70633

ID ADA70633 standard; DNA; 1344 BP.

XX

AC ADA70633;

XX

DT 20-NOV-2003 (first entry)

XX

DE Rice gene, SEQ ID 3956.

XX

KM Plant; bacterial infection; fungal infection; viral infection; rice;

XX

KM gene; ds.

XX

OS Oryza sativa.

XX

PN MO2003000898-AL.

XX

PD 03-JAN-2003.

XX

PF 22-JUN-2001; 2001WO-1B001105.

XX

PR 22-JUN-2001; 2001WO-1B001105.

XX

PA (SYN) SYNGENTA PARTICIPATIONS AG.

XX

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX

PI Katagiri F, Quan S, Tao Y, Whitcham S, Xie Z, Zhu T, Zou G;

XX

DR WPI; 2003-175290/17.

XX

PT Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to

PT bacterial, fungal or viral infection by determining or detecting plant

PT gene expression.

XX

PS Claim 6; SEQ ID NO 3956; 899bp; English.

XX

CC The present invention relates to a method (M1) for identifying genes

CC involved in plant resistance or response to pathogenic infection. M1

CC comprises identifying a gene whose expression is significantly altered in

CC the incompatible interaction of plant gene expression relative to

CC expression of the gene in an uninfected plant, in a mutant plant that

CC does not express a gene associated with response to pathogenic infection,

CC or in a corresponding incompatible or compatible interaction. (M1) is

CC useful for conferring resistance to resistance or tolerance to a plant to

CC bacterial, fungal or viral infection. The present sequence was used to

CC illustrate the invention.

XX

SQ Sequence 1344 BP; 384 A; 298 C; 331 G; 331 T; 0 U; 0 Other;

XX

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
9.7e-86	999.00	64.23%	49.23%	41.80%	1344	193	59	118	2	4

US-10-691-374-2 (1-457) x ADA70633 (1-1344)

```

QY 75 VALIIEAENVALIEUSERPHEGIALALYAGLYAASPGLYRTHRYAEPHILEA 94
DB 154 GTTGTGATGTGCGCAACATGCTCTTACGAGATGAGCAACATGATGATCAAGGCA 213

```

```

Db      1294  AATTCAGGACAGTGTCTTCACAGGCTTGCTTCC 13299
          ::: ||||| ||||| |||||
Search completed: March 4, 2005, 22:54:05
Job time : 673 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 4, 2005, 19:53:34 ; Search time 3526 Seconds

(without alignments)  
4933.457 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MWIGNSILLIIFASSIS.....VTPHCTSLSEDEALNNY 457

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delp 6.0 , Delext 7.0

Searched: 3423544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 6847908

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp

-DB=EST -QFMT=fastap -SUFFIX=rest -MISMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45

-LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=US10691374 @CGN 1.1 5180 @runat 28022005 120706 20996 -NCPU=6 -ICPU=3

-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.\*

1: gb\_esc1.\*

2: gb\_esc2.\*

3: gb\_hic.\*

4: gb\_esc3.\*

5: gb\_esc4.\*

6: gb\_esc5.\*

7: gb\_esc6.\*

8: gb\_gsa1.\*

9: gb\_gsa2.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1347	56.4	797	4	BM410796 EST585123
2	1343	56.2	786	4	BM412817 EST587144
3	1331	55.3	774	4	BM413004 EST587331
4	1292	54.1	767	4	BM412528 EST586855
5	1282	53.6	742	4	BM413282 EST587609
6	1270	53.5	749	4	BM413343 EST587670
7	1250	51.0	709	2	AM441939 EST311335
8	1212	50.7	696	2	AM442253 EST311649
9	1192	49.9	687	2	AM442335 EST311731

10	1168	48.9	677	2	AM442241	AM442241 EST311637
11	1160	48.5	675	4	BM408546	BM408546 EST582873
12	1157.5	48.4	672	2	AM221785	AM221785 EST298536
13	1128	47.2	643	2	AM223400	AM223400 EST300211
14	1118	46.8	672	2	AM222222	AM222222 EST299033
15	1110	46.4	703	4	BM536089	BM536089 EST589111
16	1109	46.4	728	4	BM413075	BM413075 EST587402
17	1107	46.3	632	2	BE334493	BE334493 EST405571
18	1107	46.3	632	2	AM442280	AM442280 EST311676
19	1106	46.3	632	2	BE332845	BE332845 EST399470
20	1102	46.1	623	4	BM409252	BM409252 EST583579
21	1099	46.0	630	2	BE331613	BE331613 EST336428
22	1097	45.9	632	4	BM410538	BM410538 EST584865
23	1094	45.8	1498	3	CNS0A7SM	BM410538 EST584865
24	1077	45.1	617	2	BE460938	BE460938 EST412357
25	1072	44.9	619	4	BM413202	BM413202 EST67529
26	1058	44.3	614	2	BE333180	BE333180 EST39709
27	1055	44.1	670	4	BM411547	BM411547 EST585874
28	1050	43.9	604	2	BE461204	BE461204 EST412623
29	1041	43.6	601	2	AM222835	AM222835 EST299646
30	1032	43.2	598	2	AM223911	AM223911 EST300722
31	1016	42.5	585	2	BE36917	BE36917 EST408035
32	1014	42.4	593	2	AM222726	AM222726 EST299537
33	1013.5	42.4	588	2	BE332533	BE332533 EST399062
34	1010	42.3	604	2	BE333321	BE333321 EST399850
35	1003	42.0	1182	9	CL972417	CL972417 OGIFFC022
36	1001	41.9	576	2	AM222864	AM222864 EST299675
37	997	41.7	1248	9	CL972416	CL972416 OGIFFC022
38	994.5	41.6	1344	9	CL972420	CL972420 OGIFFC022
39	994	41.6	572	2	AM222731	AM222731 EST299542
40	990	41.4	567	2	AM441297	AM441297 EST311593
41	989	41.4	563	2	AM441402	AM441402 EST310798
42	988	41.3	565	2	BE460910	BE460910 EST412329
43	987	41.3	567	2	AM223661	AM223661 EST300472
44	984	41.2	561	2	BE337064	BE337064 EST408182
45	981	41.0	560	2	BE332446	BE332446 EST398975

## ALIGNMENTS

RESULT 1  
BM410796  
LOCUS  
DEFINITION  
EST585123 tomato breaker fruit Lycopersicon esculentum cDNA clone  
CLES54L15 5' end, mRNA sequence.  
BM410796  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 797)  
Alcala, J., Vrebalov, J., White, R., Viston, T., Karamycheva, S.A.,  
Tsai, J., Bougari, O., Kirin, E., Utterback, T., Van Aken, S.,  
Rommig, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and  
Giovannoni, J.  
Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
Unpublished (2002)  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3  
Location/Qualifiers  
1..797  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"

```

/db_xref="taxon:4081"
/clone="cL6354L15"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_1lb="tomato breaker fruit"
/notes="Vector: pBluescriptSKmCudapt; Site_1: EcoRI;
Site_2: XhoI; supplier: Boyce Thompson Institute;
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

```

## ORIGIN

## Alignment Scores:

Pred. No.:	1.57e-144	Length:	797
Score:	1347.00	Matches:	261
Percent Similarity:	98.50%	Conservative:	1
Best Local Similarity:	98.12%	Mismatches:	3
Query Match:	56.36%	Indels:	1
DB:	4	Gaps:	0

US-10-691-374-2 (1-457) x BM410796 (1-797)

```

QY 104 SerSerArgThrProValGlnPheValValProLysAsnLysAsnTyrlLeuLeuLysGln 123
DB 2 TCATCTAGAACACCGTTCAATTGTGCTTCTAAACAGATTATCTTCTCAGCA 61
QY 124 lIeThrPheSerGlyProCyArgSerSerlIeSerValLyslIePheGlySerLeuGlu 143
DB 62 ATCACCCTTTCAAGTCCAGTCAGATCTTCTAATTCAGTAAAGATTTTGGATCTTAGAA 121
QY 144 AlaserSerylIeSerAspTyrlAspArgArgLeuTrpIleAlaPheAspSerVal 163
DB 122 GCATCTAGTAAATTTGACACTCAAGATAGAAAGCTTGCTTTGATAGTGT 181
QY 164 GlnAsnLeuValAlaGlyGlyGlyGlyThrlIeAsnGlyAsnGlyGlnValTrpTrpPro 183
DB 182 CAAATTTAGTTGGAGAGAGAGAACTATCAATGGCAATGCAAGATGTCGCCA 241
QY 184 SerSerCysAlaIleAsnLysSerLeuProCyArgAspAlaProThrAlaLeuThrPhe 203
DB 242 AGTCTTGGCAAAATTAATTAATCACTGCCATGACGAGATGCACCAAGGCTTAACTTC 301
QY 204 TrpAsnCyrlsAsnLeuLysValAsnAsnLeuLysSerLysAsnAlaGlnGlnIleHis 223
DB 302 TGGAAATTTGCAAAATTTGAAAGTAATCTAAAGATGTAAGATGCAACAAATTCAT 361
QY 224 lIeLysPheGlnSerCysThrAsnValAlaAlaserAsnLeuMetIleAsnAlaserAla 243
DB 362 ATCAATTTGAGTCAGTCATATGTTGAGCTTCAAAATTTGATATCAATGCTTCAGCA 421
QY 244 LysSerProAsnThrAspGlyValHisValSerAsnThrGlnTrpIleGlnIleSerAsp 263
DB 422 AAGAGGCCAAATACATGATGAGTCATGATCAATACCTCAATATATTCAT 481
QY 264 ThrIleIleGlyThrlYAspAspCysIleSerlIeValSerGlySerGlnAsnValGln 283
DB 482 ACTATATTGAAACGGTATGATGATTTCATTAATTTGATTCGAAATTCGAAATGTCAG 541
QY 284 AlatThrAsnIleThrCyGlyProGlyHisGlylIeSerlIeGlySerLeuGlySerGly 303
DB 542 GGCACAAATATTAATCTTGCTGCCAGTCATGATTAAGTATTTGAAAGCTTGAATCTGGA 601
QY 304 AsnSerGlnAlaTyrlValSerAsnValThrValAsnGlnAlaLyslIleIleGlyAlaGlu 323
DB 602 AATTCAGAGCTTATGTCATTAATGTTACTTAATATGAAGCCAAATTAATCGGTCGGA 661
QY 324 AsnGlyValaGlylIeLysThrTrpGlnGlyGlySerGlyGlnAlaserAsnIleLysPhe 343
DB 662 AATGAGTTAGATCAAGACTTGGCAGGAGATCTGAGCAAGTACCACTGACCATCAATTA 721

```

```

QY 344 LeuAsnValGlnUetGlnAspValLysTyrlProIleIleAspGlnAsnTyrlCys-As 363
DB 722 CTGAATTTGAAATGCAAGCGTTTGTATTCATATTAATGACCAAACTATTTGGTCA 781
QY 363 PARValGluProCys 368
DB 782 TCGAGTTTGCACATGT 797

```

## RESULT 2

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## location/Qualifiers

## 1..786

## /organism="Lycopersicon esculentum"

## /mol\_type="mRNA"

## /cultivar="TA496"

## /db\_xref="taxon:4081"

## /clone="cL6354L120"

## /tissue\_type="pericarp"

## /dev\_stage="breaker"

## /lab\_host="SOLR"

## /clone\_1lb="tomato breaker fruit"

## /notes="Vector: pBluescriptSKmCudapt; Site\_1: EcoRI; Site\_2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

## ORIGIN

## Alignment Scores:

## Pred. No.:

## Score:

## Percent Similarity:

## Best Local Similarity:

## Query Match:

Pred. No.:	4.46e-144	Length:	786
Score:	1343.00	Matches:	257
Percent Similarity:	98.09%	Conservative:	0
Best Local Similarity:	98.09%	Mismatches:	5
Query Match:	56.19%	Indels:	0
DB:	4	Gaps:	0

US-10-691-374-2 (1-457) x BM412817 (1-786)

```

QY 95 PheGluGlnAlaTrpAsnGlnAlaCysSerSerArgThrProValGlnPheValPro 114
DB 1 TTTCGCAAGCATGAAATGAAAGCATGTTCAATAGAACACCTGTTCAATTTGTGTTCT 60
QY 115 LysAsnLysAsnTyrlLeuLeuLysGlnIleThrPheSerGlyProCyArgSerSerIle 134
DB 61 AAAAAAGAAATTTATCTTCTCAAGCAATCACTTTTCAGTCCATGACATCTTCTATT 120

```

Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3.  
Location/Quality: core

FEATURES	Location/Qualifiers
source	1. .774

```

/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEG61N6"
/tissue_type="Pericarp"
/dev_stage="breaker"
/ab_tag="RC018"

```

/note="Vector: pBlueScriptSMcatadap; Site 1: EcorI; Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopersin accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

**ORIGIN**

Alignment Scores:	
Pred. No.:	1.53e-141
Score:	1321.00
Percent Similarity:	98.84%
Best Local Similarity:	98.45%
Query Match:	55.27%
DB:	4
Length:	774
Matches:	25
Conservative:	1
Mismatches:	2
Indels:	0
Gaps:	0

US-10-691-374-2 (1-457) X BM413004 (1-774)

QY	88	lysthrtyraapenllalaapheluglnalatpaenuglnalcyserseargthr	107
Db	3	AAAACATATGATAAATATTCATTTAGCAGCATGAAATGAAGCATGTTCACTGAACA	62
QY	108	ProValGlnPheValValProIlysaAnlysaenTyLeuLeuLysGlnLeuThrPheSer	127
Db	63	CCTGTTCATTTTGTCGTTCTTAAACAAGATTTCTCTCAGCAATCACCTTTCA	122
QY	128	GlyProCybargSerSerIleSerValylIlephgIySerLeuGlnIlaSerSerLys	147
Db	123	GCTTCATGCAGATCTTCTTATTCAGTAAAGATTTTGGATCCTTGAMACATCTAGTAA	182
QY	148	IleSerApTyTyLysAspArgArgLeuTrpIlealaphaAspSerValGlnIleuLeuVal	167
Db	163	ATTTCAGCTCCAAAGATAGAGGCTTGGATTCCTTGGATAGGTTTAAATTTAGTT	242
QY	168	ValGlyGlyGlyGlyThrIleasnGlyasnGlyValInvalTrpTrpProSerSerCysLys	187
Db	243	GTTGGAGAGAGGAACTATTCATGSCATGACAAAGTATGGTGCCAACTTGTCAA	302
QY	188	IleasnLysSerLeuProCysArgAspAlaPorphAlaLeuThrPheTrpAsnCysLys	207
Db	303	ATAATAATATCTCTGCATGACGGATGACCAACGGCTTAACTTCGTGAATTCAAA	362
QY	208	AsnLeuLysValAsnAsnLeuLysSerIlysaAlaIleGlnGlnIleHisIleLysPheGlu	227
Db	363	AATTYGAAGTCAATATCTTAAGGTAAATATGCAACAAATTCATTCAAATTTGAG	422
QY	228	SerCysTrpAsnValValAlaSerAsnLeuMetIleAsnAlaSerAlaLysSerProAsn	247
Db	423	TCATCCACTAATGTGTGAGCTTCAAAATTTGATGATCAATGCTTCACGAAAGGCCAAAT	482
QY	248	ThraSpGlyValHisValSerAsnThrGlnTyrlleGlnIleSerAspThrIleIleGly	267
Db	483	ACTGATGGAGTCATGATCAATATCTCAATATATTCAAATATTCGATACTCTATTATGGA	542
QY	268	ThrGlyAspAspCysIleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIle	287
Db	543	ACAGGTGATGATTTGATTTCAATGTGTCTCGATCTCAAAAGTGGACGCCCAATATAT	602

QY 288 Thrcysgllyprogllyhlsigllylleserlilegllyserleuglyserglyasnsersglua 307  
| | | | |  
Db 603 ACTTGTCGTCAGGTCATGATATGATATGAGACCTTAGATCTGGAAATTCAGAAAGCT 662  
QY 308 TyrValSerAsnValThrValAsnGluAlaLysIleIleGllyAlaGluAsnGlyValArg 327  
| | | | |  
Db 663 TATGTCGTAAATGTTACTGTAATGACCAAAATATACGTCGCCGAAATGAGCTTAGG 722  
QY 328 IleLysThr-ThrGlnGlySerGlyGlnAlaSerAsnIleLysPheLeu 344  
| | | | |  
Db 723 ATCAAGACTTTGGCAGGAGGATCTGCACAACTAGCAAAACATCAATTTCTG 774  
RESULT 4  
BM412528 767 bp mRNA linear EST 22-JAN-2002  
LOCUS EST586855 tomato breaker fruit Lycopersicon esculentum cDNA clone  
DEFINITION  
ACCESSION BM412528 1 GI:18264158  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Lycopersicon esculentum (tomato)  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 767)  
Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A.,  
Tsai,J., Bougri,O., Kirkness,E., Utechtack,T., Van Aken,S.,  
Ronnig,C.M., Fraser,C.M., Martin,G.B., Tanksey,S.D. and  
Giovannoni,J.  
Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
Unpublished (2002)  
Contact: CUGI  
TITLE  
JOURNAL  
COMMENT  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3.  
FEATURES  
source  
1..767  
Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLEG6014"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/clone\_1fb="tomato breaker fruit"  
/note="Vector: pBluescriptSkmCladap; Site 1: EcorI;  
Site 2: XhoI; supplier: Boyce Thompson Institute;  
sequencing: The Institute for Genomic Research. Fruit  
were harvested at the breaker stage (first sign of  
lycopene accumulation on the blossom end of fruit). Fruit  
were cut in half and the seeds and locules were discarded  
prior to freezing the pericarp."

Db 23 ATGCTTATCCAAAGATAGTATCTCTTCTCATTAATTTTGGCTTCATCATTTCA 82  
QY 21 ThrcysgllyserAsnValIleAspAsnLeuPheLysGlnValTyrAspAsnIleLeu 40  
| | | | |  
Db 83 ACTTGTCGTCAGGTCATGATATGATATGAGACCTTAGATCTGGAAATTCAGAAAGCT 662  
QY 41 GlnGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerLysAsnIleGlu 60  
| | | | |  
Db 143 GAACAAAGATTTGTCATGATATTTTCACTTATCTTTTATTGAGCAAAAATTTGAA 202  
QY 61 SerAsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSer 80  
| | | | |  
Db 203 AGCAACAAATTAATTTGACAGAGGTGATATAAATGAGATTAAGCATTAATGATCTTAGC 262  
QY 81 PheGlnAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGlnGlnAlaTyrAsn 100  
| | | | |  
Db 263 TTTGGAGCTAGAGGTCATGAGCAAAAATTAATTAATTTGATTTGAGCAAGCATGGAAT 322  
QY 101 GlnAlaCysSerSerArgThrProValGlnPheValAlaProLysAsnLysAsnTyrLeu 120  
| | | | |  
Db 323 GAAGCATGTCATCTAGAAACACCTGTTCAATTTGCTTCTTAAAAACAAGAAATTTATCTT 382  
QY 121 LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly 140  
| | | | |  
Db 383 CTCAGCAAAATCACCTTTTCAGTCCATGACAGATCTTCTTTTCACTMAAGATTTTGGGA 442  
QY 141 SerLeuGlnAlaSerSerLysIleSerAspTyrLysAspArgArgLeuTyrIleAlaPhe 160  
| | | | |  
Db 443 TCTTTAGAAAGCATCTAGAAATTTTTCAGATCAAAATTAAGAGCTTTGGATTTGCTTTT 502  
QY 161 AspSerValGlnAsnLeuValAlaGlyGlyGlyThrIleAsnGlyAsnGlyGlnVal 180  
| | | | |  
Db 503 GATAGTGTCAAATTTAATTTAGTTGGAGGAGGAGCACTATCAATGAGCAATGACAAAGTA 562  
QY 181 TrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaPheCotrAla 200  
| | | | |  
Db 563 TGGTGGCAAGTTTCTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 622  
QY 201 LeuThrPheThrPheCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsnAlaGln 220  
| | | | |  
Db 623 TTAACCTTCTGGAATTTGCAAAAATTTGAAAGTAATTAATTAATTAATTAATTAATTAATTA 682  
QY 221 GlnIleHisIleLysPheGlnSerCysThrAsnValValAlaSerLeuLeuMetIleAsn 240  
| | | | |  
Db 683 CAATTCATATCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 742  
QY 241 AlaSerAlaLysSerProAsnThr 248  
| | | | |  
Db 743 GCTTCAGCAAAAGAGCCCAATTAAT 766  
RESULT 5  
BM413282 742 bp mRNA linear EST 22-JAN-2002  
LOCUS EST587609 tomato breaker fruit Lycopersicon esculentum cDNA clone  
DEFINITION  
ACCESSION BM413282 1 GI:18264912  
VERSION  
KEYWORDS  
SOURCE  
Lycopersicon esculentum (tomato)  
ORGANISM  
Lycopersicon esculentum  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 742)  
Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A.,  
Tsai,J., Bougri,O., Kirkness,E., Utechtack,T., Van Aken,S.,  
Ronnig,C.M., Fraser,C.M., Martin,G.B., Tanksey,S.D. and  
Giovannoni,J.  
Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
Unpublished (2002)  
Contact: CUGI  
TITLE  
JOURNAL  
COMMENT  
Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 This clone is available through the Clemson University Genomics Institute

Seq primer: T3.

## FEATURES

source

Location/Qualifiers

1..742  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLB62H8"  
 /tissue\_type="pericarp"  
 /dev\_stage="breaker"  
 /lab\_host="SOLR"  
 /clone\_1lb="tomato breaker fruit"  
 /note="Vector: pBluescriptSMCudapt; Site 1: EcoRI; Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

## ORIGIN

## Alignment Scores:

Pred. No.: 4.6e-137 Length: 742  
 Score: 1282.00 Matches: 246  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 53.64% Indels: 0  
 DB: 4 Gaps: 0

US-10-691-374-2 (1-457) x BM413282 (1-742)

2 ValiieglnArgAsnSerIleleuLeuIleIleIlePhAlaSerSerIleSerThr 21  
 3 GTTATCCAAAGGAAATGATTTCTCTTCATTAATTTGTTGCTTCAATTCACACT 62  
 22 CybArGSerAenValiIleAphAenleuPheLygIValIlyrAphAniIleuGlu 41  
 63 TGTAGAGCAATGTTATGATGACAATTAATCAAAAGTTATGATTAATTCCTGAA 122  
 42 GlnGluPheAlaIleAphPheGlnAlaTyrlLeuSerTyrlLeuSerIleAenIleGluSer 61  
 123 CAAGAATTTGTCATGATTTTCACGTTATCTTTCTATTGAGCAAAATTAATTAAGC 182  
 62 AenAenAenIleAphLyValAenLyAenGlyIleLyValIleAenValleuSerPhe 81  
 183 AACATTAATTTGACAAAGGTTGATTAATAATGGAATTAAGTAAATTAATTAAGCTTT 242  
 82 GlyAlaLygIyAenPGLyArThrTyrrAphAenIleAlaPheGluGlnAlaTrrAenGlu 101  
 243 GGAAGCTTAAGGGTGAAGAAACATATGATTAATTAATGATTGAGCAAGCATGAGATGA 302  
 102 AlaCySerSerSerArGThrProValGlnPheValValProLyAenLyAenTyrlLeuLeu 121  
 303 GCATGTTCACTTGAACACCTGTCATATTTGGTTCCTAAACAAAGAAATTAATCTTTC 362  
 122 LygGlnIleThPhSerGlyProCybArGSerSerIleSerVallyIlePheGlySer 141  
 363 AAGCAATCACTTTTCAGTTCATGCAATCTTCTATTTCGTAATAATTTTGGATCC 422  
 142 LeuGluAlaSerSerIleSerArPyrrLyAenAryAenLyrrIleAlaPheAen 161  
 423 TTAGAGAGCATCTAGTAATAATTTCAAGCTACAAAGATAGAAAGCTTTGATCTTTGAT 482  
 162 SerValGlnAenLeuValValGlyGlyGlyTrrIleAenGlyAenGlyGlnValTrr 181  
 483 AGTGTTCAAATTTAGTGTGTGAGAGGAGAACTCAATCAAGGCAATGAGCAAGATGAG 542  
 182 TrrProSerSerCybLyIleAenLySerLeuProCybArGAspAlaProThrAlaLeu 201

DB 543 TGGCCAAATTCTTGCAAAATAAATAATCACTGCCATGAGGATGACCAACAGGCTTA 602  
 QY 202 ThrPheTrrPAsnCybLyAenleuLyValAenAenleuLySerLyAenAlaGlnGlu 221  
 DB 603 ACCTTCTGCAATTGCAAAATTTGAAAGTGAATATCTTAAGAGATTAATGCAACACA 662  
 QY 222 ILehIleLyPhePhGluSerCybThrAenValValAlaSerAenleuMetIleAenAla 241  
 DB 663 ATTCATTAATCAAAATTTGAGTCATGACCTAATGTTGATCTTCAAAATTTGATGATCAATGCT 722  
 QY 242 SerAlaLySerProAsn 247  
 DB 723 TCAGCAAGAGGCCAAAT 740

RESULT 6  
 BM413343  
 LOCUS  
 DEFINITION EST587670 tomato breaker fruit Lycopersicon esculentum cDNA clone  
 CLB63C11 5' end, mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 749)  
 Alcalá, J., Vrebalov, J., White, R., Vision, T., Karanymcheva, S.A.,  
 Teal, J., Bougri, O., Kirkness, E., Utecherback, T., Van Aken, S.,  
 Roining, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and  
 Giovannoni, J.  
 Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
 Unpublished (2002)  
 CONTACT: CUGI  
 Institution: Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 This clone is available through the Clemson University Genomics Institute  
 Seq primer: T3.

TITLE  
 JOURNAL  
 COMMENT

## FEATURES

source

Location/Qualifiers

1..749  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLB63C11"  
 /tissue\_type="pericarp"  
 /dev\_stage="breaker"  
 /lab\_host="SOLR"  
 /clone\_1lb="tomato breaker fruit"  
 /note="Vector: pBluescriptSMCudapt; Site 1: EcoRI; Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.36e-136 Length: 749  
 Score: 1278.00 Matches: 248  
 Percent Similarity: 99.60% Conservative: 0  
 Best Local Similarity: 99.60% Mismatches: 1  
 Query Match: 53.47% Indels: 1  
 DB: 4 Gaps: 0

US-10-691-374-2 (1-457) x BM413343 (1-749)

QY 74 LysValIleAenValleuSerPheGlyAlaLygIyAenGlyLyArThrTyrrAphAenIle 93

```

Db      3 AAGGTATTTATGTACTTACCTTGGAGCTTAAGGCGATGAGAAAAACATATGATATATTT
Qy      94 AlphegluGlnAlaTrpAsnGlnAlaCysSerSerArgThrProValGlnPheVala 113
Db      63 GCATTTGAGCAAGCATGAGATGAAAGCATGTTACTTGAACAACCTGTTCAATTTTGTT
Qy      114 ProLysAsnLysAsnTrpLeuLeuLysGlnIleThrPheSerGlyProCysArgSerSer 133
Db      123 CCAAAAAACAAGATTTATCTTCTCAAGCAAAATCACCTTTTCAGTCTCATGAGATCTTCT 182
Qy      134 ILeSerValLysIlePheGlySerLeuGlnAlaSerSerLysIleSerAspTrpLysAsp 153
Db      183 ATTTTCAGTAAAGATTTTGGATCTTGAAGAGCATCTAGTAAATTTTCAGACTCAAAAGAT 242
Qy      154 ATGAGTLeuTrpIleAlaPheAspSerValGlnAsnLeuValaGlyGlyGlyThr 173
Db      243 AAGAGGCTTTGGATGCTTTGATAGTGTTCAAATTTAGTTGTTGAGGAGGAGAACT 302
Qy      174 ILeAsnGlyAsnGlyGlnValTrpTrpProSerSerCysLysIleAsnLysSerLeuPro 193
Db      303 ATCAATGGCAATGACCAAGTATGCGCCAAAGTTCTTGCAAAATTAATTAATCACTGCCA 362
Qy      194 CysArgAspAlaProThrAlaLeuThrPheTrpAsnCysLysAsnLeuLysValaAsn 213
Db      363 TGCAAGGAGTGCACCAACGCGCTTAACCTTGAATTCGAAAAATTTGAAAGTGAATAT 422
Qy      214 LeuLysSerLysAsnAlaGlnGlnIleHisIleLysPheGlySerCysThrAsnVala 233
Db      423 CTAAAGGTAAATAATGCACAAATTCATATCAATTAATTAATTAATTAATTAATTAATTA 482
Qy      234 AlaSerAsnLeuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValaHisVal 253
Db      483 GCTTCAATTTGATGATCAATGCTTCAGCAAAAGGCCAAATCTATGAGATCCATGTA 542
Qy      254 SerAsnThrGlnTrpIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIle 273
Db      543 TCAATATCTCAATATATTCATTAATCTGATATCAATTAATTAATTAATTAATTAATTA 602
Qy      274 SerIleValSerGlySerGlnAsnValaGlnAlaThrAsnIleThrCysGlyProGlyHis 293
Db      603 TCAATTTGCTGATCTCAAAATGTCAGGCGCCAAATTTACTTGTGTCTCAAGTCTAT 662
Qy      294 GylIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTrpValSerAsnValThr 313
Db      663 GGTATAGATTTGAAAGCTTAGATCTGAAATTC-GAAGCTTATGTGTCTAATGTACT 721
Qy      314 ValAsnGlnAlaLysIleIleGlyAla 322
Db      722 GTAAATGAAGCCAAATTAATCGGTGCC 748

RESULT 7
AM441939 709 bp mRNA linear EST 18-MAY-2001
LOCUS EST111335 tomato fruit red ripe, TMMU Lycopersicon esculentum cDNA
DEFINITION
ACCESSION AM441939
VERSION AM441939.1 GI:6977190
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 709)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upson,J., Hansen,T., Craven,W.B., Bowman,C.L., Ahn,S.,
Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
CONTACT: CUGI
TITLE
JOURNAL
COMMENT
Clemson University Genomics Institute

```

## FEATURES

## Source

Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
Location/Qualifiers  
1..709

/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLN19E19"  
/issue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/clone\_1lb="tomato fruit red ripe, TMMU"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Giovannoni; Fruit were tagged at the  
breaker stage (first sign of lycopen accumulation on the  
blossom end of the fruit) and harvested 7 days  
post-breaker (fully red-ripe). 10 days post breaker, and  
20 days post-breaker (over-ripe). 20 day fruit which  
showed external or internal signs of pathogenesis were  
discarded. Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

## ORIGIN

## Alignment Scores:

Pred. No.: 6,31e-130 Length: 709  
Score: 1220.00 Matches: 235  
Percent Similarity: 99.58% Conservative: 0  
Best Local Similarity: 99.58% Mismatches: 1  
Query Match: 51.05% Indels: 0  
DB: 2 Gaps: 0

US-10-691-374-2 (1-457) x AM441939 (1-709)

```

Qy      219 AlaGlnGlnIleHisIleLysPheGlySerCysThrAsnValaAlaSerAsnLeuMet 238
Db      2 GCACAAACAATTCATTAATTAATTTGATGATCATGATCAATTAATTTGATCAATTTGATG 61
Qy      239 ILeAsnAlaSerAlaLysSerProAsnThrAspGlyValaHisValSerAsnThrGlnTrp 258
Db      62 ATCAATGCTTCACAAAGGCCCAATTAATCTGATGAGTCAATGATCAATTAATCTCAATAT 121
Qy      259 ILeGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly 278
Db      122 ATTCAAATATCTGATCTATTAATGACAGGATGATGATTAATTAATTAATTAATTAATTA 181
Qy      279 SerGlnAsnValaGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGly 298
Db      182 TCTCAAAATGTGAGGCCCAATTAATTAATTTACTTGTGTCTCAAGTCTGATTAATTTGGA 241
Qy      299 SerLeuGlySerGlyAsnSerGlnAlaTrpValSerAsnValThrValAsnGlnAlaLys 316
Db      242 AGCTTAGATCTGGAATTCAGAAATTCATATGCTTAATTTACTGTAATTAATTAATTAATTA 301
Qy      319 ILeIleGlyAlaGlyAsnGlyValaArgIleLysTrpTrpGlnGlnGlySerGlyGlnAla 338
Db      302 ATTATTCGTCGCCGAATTAATGAGATTAATGATCAAGACTTGGCAGGAGGATCTGCAACAGCT 361
Qy      339 SerAsnIleLysPheLeuAsnValaGlnMetGlnAspValaLysTrpProIleIleAsp 358
Db      362 ACCAACATCAAAATTTCTGAATGTGGAATGCAACACGTTAAGTATCCCATTAATTAATTAAC 421
Qy      359 GlnAsnTrpCysAspArgValaGlnProCysIleGlnGlnPheSerAlaValaGlnValaLys 378
Db      422 CAATTAATTTGATGATCGATGAAACATGTAATCAACAGTTTTCACAGCATTAAGTGA 481
Qy      379 AsnValaValaTrpGlnAsnIleLysGlyThrSerAlaThrLysValaAlaIleLysPheAsp 398
Db      482 AATGTGTGATGAGATATCAAGGCAACAGTCAACAAAGTGGCATTAATTAATTTGAT 541
Qy      399 CysSerThrAsnPheProCysGlyGlyIleIleMetGlnAsnIleAsnLeuValaGlyGln 418

```

Db 542 TGACGACAAATCTTCCATGTGTAAGGAATTAATGAGATATAATTAATTGTAAGGAA 601

Qy 419 SerGlyLyProSerGluAlaThrCysGlyValSerValHisPheAsnAsnAlaGluHisVal 438

Db 602 AGTGAAGAAACCAATCAAGGCTAGCTGCAAAAATGTCATTTTAAACATGCTGAACATGTT 661

Qy 439 ThrProHisCysThrSerLeuGluHisSerGluAspGluAlaLeuLeu 454

Db 662 AACACCACTGCACTTCACTAGAAATTTCAAGAGATGAAGCTCTTTTG 709

RESULT 8  
AM442253 696 bp mRNA linear EST 18-MAY-2001  
LOCUS EST11649 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA  
DEFINITION clone cLEN22G14 5', mRNA sequence.  
ACCESSION AM442253  
VERSION AM442253.1 GI:6977504  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 696)  
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,  
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,  
Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and  
Giovannoni,J.  
Generation of ESTs from tomato fruit tissue  
Unpublished (1999)  
CONTACT: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
FEATURES  
source  
1..696  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEN22G14"  
/cissue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/clone\_lib="tomato fruit red ripe, TAMU"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Giovannoni; Fruit were tagged at the  
breaker stage (first sign of lycopen accumulation on the  
blossom end of the fruit) and harvested 7 days  
post-breaker (fully red-ripe), 10 days post breaker, and  
20 days post-breaker (over-ripe). 20 day fruit which  
showed external or internal signs of pathogenesis were  
discarded. Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

ORIGIN  
Alignment Scores:  
Pred. No.: 5.16e-129 Length: 696  
Score: 1212.00 Matches: 231  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.71% Indels: 0  
DB: 2 Gaps: 0

US-10-691-374-2 (1-457) x AM442253 (1-696)

Qy 101 GluAlaCysSerSerGlyThrProValGlnPheValProValAsnLysAsnTyrLeu 120

Db 3 GAAGCAATGTCATCTGAAACACCTGTTCAATTTGGTTCTAAAACAAAGATTAATCTT 62

Qy 121 LeuLyGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly 140

Db 63 CTCAGCAAAATCACCTTTTCAGGTCCATGACAGATCTTATTTTCAGTAAGATTTTGGAA 122

Qy 141 SerLeuGluAlaSerSerIleIleSerAspTyrLysAspAspArgLeuTrpIleAspLe 160

Db 123 TCCCTTAAGAAAGCATCTAGTAATAATTTTCAGCTVCAAGAAAGAGCTTTGGATGCTTTT 182

Qy 161 AspSerValGlnAsnLeuValAlaGlyGlyGlyThrIleAsnGlyAsnGlyVal 180

Db 183 GATAGGTTCAAAATTTTACTGTTTGAGAGAGAGAGAACTATCATATGCAATGAGCAAGTA 242

Qy 181 TrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaProThrAla 200

Db 243 TGGTGCCAAATGCTCCGCAAAATTAATCACTGCGCATGAGGAGATCACCACAGCGCC 302

Qy 201 LeuThrPheThrAspCysGlyValSerLeuValAsnAsnLeuLysSerIleAsnAlaGln 220

Db 303 TTAACTCTTGGAATTCGCAAAATTTGAAAGTGAATTAATCTAAAGAGTAAAAATGACAA 362

Qy 221 GlnIleHisIleLysPheGluSerCysThrAsnValAlaAlaSerAsnLeuMetIleAsn 240

Db 363 CAATTCATATCAATTAATTTAGTCATGCACTAATGTTGAGCTTCAAAATTTGATGATCAT 422

Qy 241 AlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260

Db 423 GCTTCAGCAAAAGCCCAAAATCTGATGAGATGTCATGTAATCAATTAATTAATTTCAA 482

Qy 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280

Db 483 ATATCTGATATCATTAATTTGGAACAGTGATGATGATTAATTTCAATGTTTCTGATCTCA 542

Qy 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300

Db 543 AATGTCAGAGCCCAAAATTTACTTGTCGCCAGGTCAGATGATTAAGTTTGGAACTTA 602

Qy 301 GlySerGlyAsnSerGluAlaTyrValSerAsnValThrValAsnGluAlaLysIleIle 320

Db 603 GGATCTGAAATTCAGAACTTATGCTATGCTAATGTTACTGTAATGAACCAAAATTAATC 662

Qy 321 GlyAlaGluAsnGlyValArgIleLysThrTrp 331

Db 663 GGATGCCGAAATGAGATGATCAAGACTTGG 695

RESULT 9  
AM442335 687 bp mRNA linear EST 18-MAY-2001  
LOCUS EST111731 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA  
DEFINITION clone cLEN22L13 5', mRNA sequence.  
ACCESSION AM442335  
VERSION AM442335.1 GI:6977586  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 687)  
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,  
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,  
Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and  
Giovannoni,J.  
Generation of ESTs from tomato fruit tissue  
Unpublished (1999)  
CONTACT: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
FEATURES  
source  
1..687  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"

/db xref="taxon:4081"  
/clone="clEN2213"  
/tissue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/clone\_lib="tomato fruit red ripe, TAMU"  
/note="Vector: pBluescript SK(-); site\_1: EcoRI, site\_2:  
XhoI; supplier: Giovannoni; Fruit were tagged at the  
breaker stage (first sign of lycopene accumulation on the  
blossom end of the fruit) and harvested 7 days  
post-breaker (fully red-ripe), 10 days post breaker, and  
20 days post-breaker (over-ripe). 20 day fruit which  
showed external or internal signs of pathogenesis were  
discarded. Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

## ORIGIN

## Alignment Scores:

Pred. No.:	1 04e-126	Length:	687
Score:	1192.00	Matches:	226
Percent Similarity:	99.56%	Conservative:	1
Best Local Similarity:	99.12%	Mismatches:	1
Query Match:	49.87%	Indels:	0
DB:	2	Gaps:	0

US-10-691-374-2 (1-457) x AM442335 (1-687)

QY 76 llaaenvalleuSerPheGlyAlaYseGlyAspGlyThrTyraAspAnlleAlaPhe 95  
DB 3 ATTATGTACTTAGCTTTGAAAGCTTAGGCTATGGAAGAAACATATGAAATATGCAATTT 62  
QY 96 GluGlnAlaTPaenGluAlaCyseSerSerArgThrProValGlnPheValProlys 115  
DB 63 GAGCAAGATGGAATGGAAGCATGTTCATAGAAACACCTTCATTTGTTGTTCCAA 122  
QY 116 AsnLysAsnTyrlLeuLeuLysGlnlleThrPheSerGlyProCysArgSerSerlleSer 135  
DB 123 AACAGAAATATCTTCCAGCAAAATCACTTTACAGTTCACAGATCTTCTATTTCA 182  
QY 136 ValLysllePheGlySerleuGluAlaSerSerLyslleSerAspTyrlLysAspArg 155  
DB 183 GTTAAAGATTTTGGATCTTAGAAGCATCTGTAATTAATTTCACTACAAAGATGAAAG 242  
QY 156 LeuTrpIleAlaPheAspSerValGlnAsnLeuValGlyGlyGlyThrIleAsn 175  
DB 243 CTTCGATGCTTTGATGATGTTCAAAATTTTGTTGGAGAGAGAGAACTATCAT 302  
QY 176 GlyAsnGlyGlnValTrpTrpProSerSerCysLyslleAsnLysSerleuProCysArg 195  
DB 303 GGCATATGACAAATGATGTGGCCAAAGTCTTGCAAAATTAATTAATCACTGCATGCAG 362  
QY 196 AspAlaProThrAlaLeuThrPheTrpAsnCylyAsnLeuLysValAsnAsnLeuLys 215  
DB 363 GATGACCAACGGCTTAACTTCTGGAATTCAGAAATTTGAAAGTCAATATATTAAG 422  
QY 216 SerLysAsnAlaGlnGlnlleHsiLysPheGlySerCysThAsnValValAlaSer 235  
DB 423 AGTAAAGAAAGCAACAAATTCATATCAATTTGATGATGACCTAACTTGTGACTTCA 482  
QY 236 AsnLeuMetlleAsnAlaSerAlaLysSerProAsnTrpAspGlyValHsiValSerAsn 255  
DB 483 AATTGATGATCAATGCTTCAAGCAAGAGCCCAAAATTCATATGAGATCAATATTAAT 542  
QY 256 ThrGlnTyrlleGlnlleSerAspThrIlelleGlyThrGlyAspAspCyslleSerlle 275  
DB 543 ACTCAATATATCAATATCTGATCTATTTATGGAAGAGTGATGATGATTTCAAT 602  
QY 276 ValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHsiGlylle 295  
DB 603 GTTTCGATCTCAAAATGTCAGGCCACACATATTTCTTGTGTCCAGGTCATGTGATA 662  
QY 296 SerlleGlySerleuGlySerGly 303  
DB 663 AGTATGGAAGCTTAGATCTGA 686

## RESULT 10

AM442241

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 677  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="74496"  
/db\_xref="taxon:4081"  
/clone="clEN22C16"  
/tissue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/clone\_lib="tomato fruit red ripe, TAMU"  
/note="Vector: pBluescript SK(-); site\_1: EcoRI, site\_2:  
XhoI; supplier: Giovannoni; Fruit were tagged at the  
breaker stage (first sign of lycopene accumulation on the  
blossom end of the fruit) and harvested 7 days  
post-breaker (fully red-ripe), 10 days post breaker, and  
20 days post-breaker (over-ripe). 20 day fruit which  
showed external or internal signs of pathogenesis were  
discarded. Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

## ORIGIN

## Alignment Scores:

Pred. No.:	6 04e-124	Length:	677
Score:	1168.00	Matches:	224
Percent Similarity:	99.56%	Conservative:	0
Best Local Similarity:	99.56%	Mismatches:	1
Query Match:	48.87%	Indels:	0
DB:	2	Gaps:	0

US-10-691-374-2 (1-457) x AM442241 (1-677)

QY 189 AsnLysSerleuProCysArgAspAlaProThrAlaLeuThrPheTrpAsnCylyAsn 208  
DB 1 AATTAATCACTGCAATGAGGATGCAACACGGCTTAACTTCTGGAATTCAGAAAT 60  
QY 209 LeuLysValAsnAsnLeuLysSerLysAsnAlaGlnGlnlleHsiLysPheGlySer 228  
DB 61 TTGAAGAGGAAATTAATCTTAAAGAGTAAATGCAACAAATTCATATCAATTTGAGTCA 120  
QY 229 CysThrAsnValValAlaSerAsnLeuMetlleAsnAlaSerAlaLysSerProAsnThr 248  
DB 121 TGCATTAATGTTGAGCTTCAAAATTTGATGATCAATCTTCAGCAAGAGCCCAATACT 180  
QY 249 AspGlyValHsiValSerAsnThrGlnTyrlleGlnlleSerAspThrIlelleGlyThr 268  
DB 181 GATGAGTCCATGATCAATCAATCTCAATATTAATCTGATATCTATTTATGAGACA 240



QY 269 GYAAPPAPCYAIIeserIleValSerGIySerGlnAsnValGlnAlaThrAsnIleThr 288  
DB 241 GGATGATGATGATTAATTCATTTGTTCTGATCTCAAAATGTGCGAGCCCAATATATTACT 300  
QY 289 CYSGIYProGIYHISGIYIleSerIleGIySerLeuGIySerGIYAsnSerGIuAlaTYr 308  
DB 301 TGTGCTCCAGCTCATGCTATTAATGATTTGGAAGCTTGATTCGGAATTCAGAGCTTAT 360  
QY 309 ValSerAsnValThrValAsnGlnAlaValIleIleGIYAlaGlnAsnGIYValArgIle 328  
DB 361 GGTCTAAATGTTACTTAATATGAAAGCAAAATATATCGTGCCGAAATATGAGTTAGATTC 420  
QY 329 IYerThrTTGInGIYGIYSerGIYGIYAlaSerAsnIleIlePheLeuAsnValGIUmec 348  
DB 421 AAGACTTGGCAGGAGAGATCTGACAAAGTCAGCAACATCAAAATTTCTGATGTGGAAG 480  
QY 349 GlnAEPValIYerTYrProIleIleIleAEPGlnAsnTYrCYeAPArgValGIUProCYs 368  
DB 481 CAAGAGCTTAAGTATCCATTAATTAAGACCAAAATATTTGATGATGATGAAACATGT 540  
QY 369 IIEGInGlnPheSerGlnAlaValGlnValIYAsnValValTYrGIUAsnIleIYSGIYThr 388  
DB 541 ATACAAACAGTTTTCACAGCAGTTCAAGTGAAGAAATGTGTATGAGATATCAAGGCGACA 600  
QY 389 SerAlaThrIYValAlaIleIYAsnPheAPCYeSerThrAsnPheProCYSGIUGIYIle 408  
DB 601 AGTGCACAAAGAGTGCCATTAATTTGATGTGACGACAAACCTTTCATGTGAAAGATT 660  
QY 409 IIEmerGIUAsnIle 413  
DB 661 ATATGAGAAATTA 675

RESULT 11  
BM408546 675 bp mRNA linear EST 22-JAN-2002  
LOCUS EST82873 tomato breaker fruit Lycopersicon esculentum cDNA clone  
DEFINITION CLEB45C7 5' end, mRNA sequence.  
ACCESSION BM408546  
VERSION BM408546.1 GI:18260176  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 675)  
Alcala,J., Vrebalov,J., White,R., Vision,T., Karanycheva,S.A.,  
Roming,C.M., Frazer,C.M., Kirkness,E., Utterback,T., Van Aken,S.,  
Giovannoni,J.  
Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
Unpublished (2002)  
CONTACT: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3.

FEATURES  
source  
Location/Qualifiers  
1..675  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLEB45C7"  
/ribose\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/clone\_lib="tomato breaker fruit"  
/note="Vector: pBluescriptskmCudapc; Site\_1: EcoRI;

Site 2: XhoI; supplier: Boyce Thompson Institute;  
sequencing: The Institute for Genomic Research. Fruit  
were harvested at the breaker stage (first sign of  
lycopene accumulation on the blossom end of fruit). Fruit  
were cut in half and the seeds and locules were discarded  
prior to freezing the pericarp."

ORIGIN  
Alignment Scores:  
Pred. No.: 5,06e-123 Length: 675  
Score: 1160.00 Matches: 222  
Percent Similarity: 99.11% Conservative: 0  
Best Local Similarity: 99.11% Mismatches: 2  
Query Match: 48.54% Indels: 0  
DB: 4 Gaps: 0

US-10-691-374-2 (1-457) x BM408546 (1-675)

QY 104 SerSerArgThrProValGlnPheValIleProLYAsnLYAsnTYrLeuLYSGIn 123  
DB 2 TCATCTAGAACCTGTTCAATTTGTGTCTTAAACAAATTAATCTTCAGACAA 61  
QY 124 IIErnpheSerGIYProCYsArgSerIleSerValIYIlePheGIYSerLeuGIU 143  
DB 62 ATCACCCTTTTCAGTCCATGACATCTTCTATTTAGTAAAGATTTTGGATCCTTAGAA 121  
QY 144 AlAserSerIYIleSerAPTYrIYAsnPAArgLeuTYrIleAlaPheAsnSerVal 163  
DB 122 GCATCTAGAAATTTTCACAGCTACAAAGATTAAGCTTTGATGCTTTTAATGATGTT 181  
QY 164 GlnAEPValIValGIYGIYIYIleAsnGIYAsnGIYQlnValTYrTPPro 183  
DB 182 CAAATTTAGTGTGTGAGAGAGAGAACTTAATGCAATGCAATGCAATGATGTGGCA 241  
QY 184 SerSerCYeLYIleAsnLYSerLeuProCYsArgAspAlaProThrAlaLeuThrPhe 203  
DB 242 ACTTCTGCAAAATTAATTAATCACTGCCATGCAAGGATGCAACAGGCTTAACCTTC 301  
QY 204 TTPAsnCYeLYAsnLeuLYValAsnAsnLeuLYSerLYAsnAlaGlnIleHis 223  
DB 302 TCGAATTCGCAAAATTTGAAAGTGAATTAATCAAGATAAGTCAACAAATTCAT 361  
QY 224 IIEYAsnPheGlnSerCYerThrAsnValIleAlaSerAsnLeuMerIleAsnAlaSerAla 243  
DB 362 ATCAATATTGAATCAAGCACTAATGTGTAGCTTCAAAATTAATGATGATGATGATGCA 421  
QY 244 IYSerProAsnThrAsnGIYValIleValSerAsnThrGlnTYrIleGlnIleSerAsp 263  
DB 422 AAGAGCCCAATATGATGATGATCCATGATCAATATCAATATATCTGAT 481  
QY 264 ThrIleIleGIYThrGIYAsnAPCYeIleSerIleValSerGIYSerGlnAsnValGln 283  
DB 482 ACTATATTGGAACAGTGATGATTTCAATTTGATTTCTCGAATCTCAAAATGTGCGAG 541  
QY 284 AlAThrAsnIleThrCYSGIYProGIYHISGIYIleSerIleGIYSerLeuGIYSerGIY 303  
DB 542 GGCACAAATATTAATCTGTGCTCAGGTCAAGTCAAGTATTAAGTATTAAGAGCTTGAATCGAA 601  
QY 304 AsnSerGIuAlaTYrValSerAsnValThrValAsnGlnAlaValIleIleGIYAlaGln 323  
DB 602 AATTCAGAAAGCTTAATGTGTCTAATGATCTGTAATTAAGAACCAAAATATTCGGTGGCGA 661  
QY 324 AsnGIYValArg 327  
DB 662 AATGAGATTAGG 673

RESULT 12  
AW221785 672 bp mRNA linear EST 18-MAY-2001  
LOCUS EST98596 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA  
DEFINITION clone CLENI122, mRNA sequence.  
ACCESSION AW221785  
VERSION AW221785.1 GI:6533469

KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 672)  
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Rønning,C.M., Frazer,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
TITLE Generation of ESTs from tomato fruit tissue  
JOURNAL Unpublished (1999)  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source  
1..672  
Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA96"  
/db\_xref="taxon:4081"  
/clone="CLEN11F14"  
/tissue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/clone\_lib="tomato fruit red ripe, TAMU"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Giovannoni; Fruit were tagged at the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe). 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

## ORIGIN

Alignment Scores:  
Pred. No.: 9.77e-123 Length: 672  
Score: 1157.50 Matches: 221  
Percent Similarity: 99.10% Conservative: 0  
Best Local Similarity: 99.10% Mismatches: 1  
Query Match: 48.43% Indels: 1  
DB: 2 Gaps: 1

US-10-691-374-2 (1-457) x AW221785 (1-672)

QY 90 TyrAspAsnIleAlaPheGluGlnAlaTPAsnGluAlaCysSerSerArgThrProVal 109  
Db 3 TATGATATATTTGCAATTTAGCAAGCATGGAAATGAAGCATGTTCACTTGAACACCTGTT 62  
QY 110 GluPheValValProLysAsnLysAsnTYrLeuLeuLysGlnIleThrPheSerGlyPro 129  
Db 63 CAATTGTGTGTTCTTAATAAACAAGATTTCTTCTCAACAATTAACCTTTTCAGGTCA 122  
QY 130 CysArgSerSerIleSerValLysIlePheGlySerLeuGluAlaSerSerLysIleSer 149  
Db 123 TGGAGATCTTCTATTTCACTTAAGATTTTGGATCTTGAAGCATCTAGTAAATTTCA 182  
QY 150 AspTYrLysAspArgLysLeuTrpIleAlaPheAspSerValGlnAsnLeuValGly 169  
Db 183 GACTACAAAGATAGAAGCTTTGGATGCTTTGATAGTGTTCAAATTTAGTTGGA 242  
QY 170 GlyGlyGlyThrIleAsnGlyAsnGlyGlnValTrpProSerSerCysLysIleAsn 189  
Db 243 GAGGAGGAACATATCAATGCGCAATGACAAAGATGTGGCAAGTTCTTGCAAAATTAAT 302  
QY 190 LysSerLeuProCysArgAspAlaProThrAlaLeuThrPheTrpAsnGlyLysAsnLeu 209  
Db 303 AAATCATGCGCATGCAAGGATGCAACAAAGCGCTTTAACTTCTGAAATTCAGAAATTTG 362

QY 210 LysValAsnAsnLeuLysSerLysAsnAlaGlnGlnIleHisIleLysPheGluSerCys 229  
Db 363 AAAGTGATATATCTAAAGAGTAATAATGCAACAACAAATTCATATCAATTTAGTCATCC 422  
QY 230 ThrAsnValValAlaSerAsnLeuMetIleAsnAla---SerAlaLysSerProAsnThr 248  
Db 423 ACTAATCTGTACTCTCAATTTGATGATCAATGCTTTTTCAGCAAAAGGCCCAATTAAT 482  
QY 249 AspGlyValHisValSerAsnThrGlnTYrIleGlnIleSerAspThrIleIleGlyThr 268  
Db 483 GATGAGTCATGATATCAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 542  
QY 269 GlyAspAspCysIleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThr 288  
Db 543 GGTGATGATTTGATTTTCAATTTGATGATCAATGCTTTTTCAGCAAAAGGCCCAATTAAT 602  
QY 289 CysGlyProGlyHisGlyIleSerIleGlySerLeuGlySerGlyLysSerGluAlaTYr 308  
Db 603 TGTGTCAGTCATGATATAGTATTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTAT 662  
QY 309 ValSerAsn 311  
Db 663 GTGTCTTAAT 671

RESULT 13  
AM223400 643 bp mRNA linear EST 18-MAY-2001  
LOCUS EST300211 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA  
DEFINITION clone CLEN11F14, mRNA sequence.  
ACCESSION AM223400  
VERSION AM223400.1 GI:535084  
KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 643)

REFERENCE Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Rønning,C.M., Frazer,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
TITLE Generation of ESTs from tomato fruit tissue  
JOURNAL Unpublished (1999)  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source  
1..643  
Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA96"  
/db\_xref="taxon:4081"  
/clone="CLEN11F14"  
/tissue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/clone\_lib="tomato fruit red ripe, TAMU"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Giovannoni; Fruit were tagged at the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe). 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

ORIGIN  
Alignment Scores:

Pred. No.:	2.35e-119	Length:	644
Score:	1128.00	Matches:	213
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	47.20%	Indels:	0
DB:	2	Gaps:	0

US-10-691-374-2 (1-457) X AW223400 (1-643)

OY	90	YrAaPaAnIleAlaPheGIuGIInIaTPaNGuAlaCySeSeSeXrGrTProVal	109
Db	3	TAAGTAATATTGCAATTTAGCAGACGATGGAATGAGCATGTTCACTTGAGACCTGTT	62
OY	110	GInPheValProLYsaNLYsaenTYrLeuLeuLYaGInIleThrPheSerGIPro	129
Db	63	CAATTGGTGGTCCATAAAACAAGAATYATCTTCCAAACCAATCACCTTTTCAGGTCCA	122
OY	130	CydaGSeSeSerIleSeValLYsIlePheGIySeLeuGInIaSeSerLYsIleSer	149
Db	123	TGCAATCTTCATATTCAGTAAGAATTTTGGATCTTGAAGACATCTGATMAAATTTCA	182
OY	150	AsPYrLYsaPaArGArGLeuTrpIleAlaPheAsPSeValGInaSnLYeValValIGLY	169
Db	183	GACTCAAAAGATAGAAAGCCTTTGGANTGCTTTTGARAGGTTCMAAATTTAGTGTGGA	242
OY	170	GIyGIyGIyThrIleaNGLyaNGIyGInValTTPTrPProSeSeCySeLYsIleAen	189
Db	243	GGAGAGGAAACTATGCAATGGACAAGTAGTGTGGCCAAATTCCTTGCAAAATAAAT	302
OY	190	LYdSeSerLeuProCYaArGAsPaAlaProThrAlaLeuThrPheTrPaSnCYeLYsaNLeu	209
Db	303	AAATCACTGCATGCAGGAGATCACCAAGGCCCTTAACCTTGGAATTCGCAAAAATTTGG	362
OY	210	LYdValaAsnAsnLeuLYSeSerLYsaNlaGInGInIleHsIleLYsaPheGIuSeCYs	229
Db	363	AAAGGAAATTAATCTMAAGGTAATAATGCACAACAATTCATATCAAAATTTGATGATGC	422
OY	230	ThraenValValaISeSerAsnLeuMetIleAsnAlaSeSerAlaLYSeSerProAsnThraP	249
Db	423	ACTAATGTTGTAAGCTTCAAAATTTGATGATCAATGCTTCAGCAAAAGGCCAAATATCTGAT	482
OY	250	GIyValHsIleValSeSerAsnThrGInTYrIleGInIleSeSerAsPThrIleIleGIyThrGIy	269
Db	483	GGAGTCCATGTAATCTCAATATCTCAATATATCTCAATATCTGATATTTAGGAACAGGT	542
OY	270	AsPaAPCYaIleSeSerIleValSeSerGIySeSerGInaSnValGInaIaThrAsnIleThrCYs	289
Db	543	GATGATTTGATTTCAATTTGTTCTGGATCTCAAAAAGTGCAGGCCACIAATATTAATCTGT	602
OY	290	GIyProGIyHsIleGIyIleSeSerIleGIySeSerLeuGIySeSer	302
Db	603	GGTCCAGGTCAAGTATTAAGTATTTGGAAAGCTTAAGGATCT 641	

RESULT	14
LOCUS	AM222222
DEFINITION	672 bp mRNA linear EST 18-MAY-2001 EST229903 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
ACCESSION	AM222222
VERSION	AM222222.1 GI:6533906
KEYWORDS	EST.
SOURCE	Lycopersicon esculentum (tomato)
ORGANISM	Lycopersicon esculentum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 672)
REFERENCE	Alcala,J., Vrebalov,J., White,R., Matern,A.L., Helt,I.B., Liang,F., Upchurch,J., Hansen,T., Craven,M.B., Bowman,C.B., Ahn,S., Kominig,A.C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
AUTHORS	Generation of ESTs from tomato fruit tissue Unpublished (1999)
TITLE	
JOURNAL	

**COMMENT**

**Contact:** CMGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence. *Genetics* 161:103-110 (2002)

## FEATURE

```

"organism"="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TM496"
/db_xref="caxon:4081"
/clone="cLEN716"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_1ib="tomato fruit red ripe, TMU"
/note="vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
XhoI, supplier: Giovannoni; Fruit were tagged at the
breaker stage (first sign of lycopene accumulation on the
blossom end of the fruit) and harvested 7 days
post-breaker (fully red-ripe), 10 days post breaker, and
20 days post-breaker (over-ripe). 20 day fruit which
showed external or internal signs of pathogenesis were
discarded. Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

```

## ORIGIN

### Alignment Scores:

pred. No.:	3.6e-118	Length:	67
Score:	1118.00	Matches:	21
Percent Similarity:	99.54%	Conservative:	0
Best Local Similarity:	99.54%	Mismatches:	0
Query Match:	46.78%	Indels:	1
DB:	2	Gaps:	0

US-10-691-374-2 (1-457) X AW222222 (1-672)

Oy	1	MevallileglnarygsnserlleuleuenuellelephalaSerSerIeser	20
Dd	24	AAGGTATTCGAAAGGAATAGTATTCTCCTTCATTAATTATTTGGCTTCATCA	83
Oy	21	ThrCyatgSeranValilleasphaenleupheylginValTyraSpsnnlleu	40
Dd	84	ACTTGAGAAGCAATTTAATGATGCATTTATTCAAACAAGTTAATAATATCTT	14
Oy	41	GluGlulnPhelalhisaphegnlalytleusertyrleuSerlySaasnillegu	60
Dd	144	GAACAGAAATTCCTATATATTTCAAGCTTAATCTTCTTAATTAGCAAAAATATGAA	20
Oy	61	SerAsnaSnasnilleaplysaValaplysancilylelasyValleasnValleuser	80
Dd	204	AGCAACATATATTTGACAAAGGTGATMAAAATGGATTAAGAATTAATGACTTAGC	26
Oy	81	PheglValalyselyspeglytyrthrTyraSpsnnllealapheglulnlaatrpaSn	100
Dd	264	TTTGGAGCTAAAGGTGATGAAAAACATATGATAATATGCAATTTGACAAAGCATGAT	32
Oy	101	GluAlaCyseSerSerzThrProvalGlnPheValProLySaSnlySaasnTyrlieu	120
Dd	324	GAAGCATGTTCACTGAACACCTGTCAATTTGGTTCCATAAAAACAAGATTATCTT	38
Oy	121	LeuLyeGlnlIethrPheSerGlyProCyatgSerSerIleservallyellPhegly	140
Dd	384	CTAACGAAATCACCTTTCAAGTCCATCGCAATCTTCTAATTCAGTAAAGATTTTTGGA	44
Oy	141	SerleuGlnlaseSerlyslleSerapyrlylsasparGargyleutripallealphe	160
Dd	444	TCTTGAGAACACTGATAAATTTCAAGCTCAAAAGATAGAAGCTTTGGATTTGCTTT	50
Oy	161	AhpServaGlnaSnleuValalgylyglytyrllleanglySaSnlyQlnVal	180
Dd	504	GATAGGTCTCAAAATTTAGTTGTGAGAGAGAGAACATTAACAAGCAATGACCAAGTA	56

Qy	41	GlugInGluPheAlaHisapPheGlnAlaTyrLeuSerTyrLeuSerLysAsnIleGlu	60
Db	166	GAACAAGATTGGCTCAAGATTTCACAGCTTACTTCTTATTATTGGACAAAATTTCGA	22
Qy	61	SerAsnaAsnaniIeapLysValAspLysAsnGlyIleLysValIleAsnValLeuSer	80
Db	225	AGCAACATATATATTGGACAAGTTGATTAATAATGGATTAAAGTATTAATCTACTTGC	28
Qy	81	PheGlyAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGluGlnAlaTyrAsn	100
Db	285	TTTGACCTAAAGGTGATGAGAAACATATGATTAATTTGCAATTTGGACCAAGCATGGAA	34
Qy	101	GluAlaCysSerSerArgThrProValGlnPheValProLysAsnLysAsnTyrLeu	120
Db	345	GAACACATGTCATCTGAAACACCTGTTCAATTTGGTTCTTCCTTAATAAACAAATTA	40
Qy	121	LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly	140
Db	405	CTCAGCAATTCACCTTTCAAGTCACAGCAAGATCTTCAATTTACAGTAAAGATTTT	46
Qy	141	SerLeuGluAlaSerSerLysIleSerArgTyrLysAspArgLysLeuTyrIleAlaPhe	160
Db	465	TCCTTGAAAGCATCTGTAAATTTCAAGCTACAAAGATTAAGAAGCTTTGAATGCTTT	52
Qy	160	eapSerValGlnAsnLeuValIleGlyGlyGlyGlyThrIleAsnGlyAsnGlyGln	180
Db	525	TGATAGCTTCAAAATTTAGTTGTTGGAGAGAGAGAAATCAATCAATGGCAATGGACA	58
Qy	180	AlaTyrTyrProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaProThr	200
Db	585	TATGGGGGCAAGTTCTGGCAAAATAAATAATCACTGCCATGAGAGGAGCAACCA	64
Qy	200	IaLeuThrPheTyrAsnCysLysAsnLeuLysValIleAsnLeuLysSerLys	217
Db	645	CCTTAACCTTCGTAATGCAAAAATTTGAAAGTAAATCTTAAGAAGTAA	697

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 5, 2005, 00:17:31 ; Search time 661 Seconds  
(without alignments)

4102.523 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 2 MYQNSILLIIRASIS.....VTPHCTSLISDEALNNY 457

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1401638 seqs, 2966923429 residues

Total number of Hits satisfying chosen parameters: 10803276

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xlp  
-O=/cg2n\_1/USPTO.spool.p/US10691374/runcat\_28022005\_120708\_21090/app\_query.fasta.1.647  
-DB=Published Applications NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10691374.0CGN.1.1.723 @runcat\_28022005\_120708\_21090  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOPE=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cg2n\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2n\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cg2n\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cg2n\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cg2n\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cg2n\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cg2n\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cg2n\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cg2n\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cg2n\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cg2n\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cg2n\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cg2n\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cg2n\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cg2n\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cg2n\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cg2n\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cg2n\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cg2n\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
20: /cg2n\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
21: /cg2n\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
22: /cg2n\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1820	76.2	5822	10	US-09-924-197-1
2	1506	63.0	7456	18	US-10-691-374-1
3	1498	62.7	7456	18	US-10-691-374-47
4	1498	62.7	7456	18	US-10-691-374-49
5	1127	47.2	1621	18	US-10-739-930-530
6	1113	46.6	1631	13	US-10-151-668-1
7	1103	46.2	1657	18	US-10-787-958-31
8	1103	46.2	1707	18	US-10-739-930-1548
9	1011	42.3	1182	18	US-10-437-963-24608
10	1010	42.3	1359	18	US-10-437-963-24608
11	988.5	41.4	1335	18	US-10-437-963-28561
12	971.5	40.6	1168	17	US-10-260-238-618
13	940	39.3	1372	18	US-10-425-115-33752
14	870	36.4	1308	17	US-10-425-114-30261
15	808	33.8	1573	18	US-10-437-963-52002
16	785	32.8	1886	18	US-10-437-963-31774
17	770	32.2	1623	18	US-10-425-115-82276
18	739	30.9	1723	18	US-10-437-963-73025
19	734	30.7	1512	18	US-10-437-963-70206
20	733.5	30.7	1822	18	US-10-739-930-2789
21	728	30.5	1631	17	US-10-425-114-9445
22	728	30.5	1984	17	US-10-425-559-92601
23	727	30.4	1673	17	US-10-362-091-3
24	726.5	30.4	2007	18	US-10-425-115-3138
25	723.5	30.3	1479	17	US-10-425-599-141807
26	709.5	29.7	1733	18	US-10-739-930-1144
27	702.5	29.4	1185	9	US-09-938-842A-2520
28	702.5	29.4	1185	11	US-09-938-842A-2520
29	701.5	29.4	1203	18	US-10-437-963-52557
30	700	29.3	1884	17	US-10-425-114-3023
31	700	29.3	1919	18	US-10-425-115-129068
32	699	29.2	1633	17	US-10-425-114-26774
33	691.5	28.9	1844	18	US-10-739-930-2788
34	682	28.5	2219	18	US-10-437-963-55617
35	681	28.5	2174	18	US-10-437-963-59867
36	679.5	28.4	1374	18	US-10-437-963-43683
37	666	27.9	1182	18	US-10-437-963-13631
38	660	27.6	1599	17	US-10-424-599-33478
39	656.5	27.5	1943	18	US-10-425-115-10070
40	655	27.4	1388	17	US-10-425-114-4296
41	655	27.4	1473	17	US-10-425-114-17989
42	654.5	27.4	1455	17	US-10-425-114-14146
43	654.5	27.4	1472	17	US-10-424-599-95670
44	654.5	27.4	1701	18	US-10-425-115-44782
45	653.5	27.3	1312	17	US-10-425-114-14060

#### ALIGNMENTS

RESULT 1  
US-09-924-197-1  
Sequence 1, Application US/09924197  
Publication No. US20030018993A1  
GENERAL INFORMATION:  
APPLICANT: Guttererson, Neal  
TITLE OF INVENTION: Improved Methods of Gene Silencing Using Inverted  
FILE REFERENCE: 012176-01081005  
CURRENT FILING DATE: 2001-08-07  
PRIOR APPLICATION NUMBER: 60/225,508  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 5822  
TYPE: DNA  
ORGANISM: Agrobacterium tumefaciens

US-09-924-197-1

## Alignment Scores:

Pred. No.:	4,28e-200	Length:	5822
Score:	1820.00	Matches:	345
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.71%	Mismatches:	0
Query Match:	76.15%	Indels:	0
DB:	10	Gaps:	0

US-10-691-374-2 (1-457) x US-09-924-197-1 (1-5822)

```

QY 112 ValValProLysAsnLysAsnLysLeuLysGlnIleThrPheSerGlyProCysArg 121
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1387 ATGTTCTTAAACCAAGAAATTAATCTTCTCAGCAAAATCACCTTTTCAAGTCCATGCA 1446
QY 132 SerSerIleSerValLysIlePheGlySerLeuGlnLysSerLysIleSerAspTyr 151
   TCTTCTATTTTCAGTAAAGATTTTGGATCTTAGAACATCTAGTAAATTTTCAGACTAC 1506
Db 1447 TCTTCTATTTTCAGTAAAGATTTTGGATCTTAGAACATCTAGTAAATTTTCAGACTAC 1506
QY 152 LysAspArgArgLeuTrrIleAlaPheAspSerValGlnIleAsnLeuValValGlyGly 171
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1507 AAGACATGAAGGCTTTGATTTGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1566
QY 172 GlyThrIleAsnGlyAsnGlyGlnValTrrTrrProSerSerCysLysIleAsnLysSer 191
   1567 GGAACATCATGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626
Db 1567 GGAACATCATGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626
QY 192 LeuProCysArgArgAspAlaPrrThrAlaLeuThrPheTrrAsnCysLysAsnLeuLysVal 211
   1627 CTGCATGACGAGGATGACACCAAGGCTTAACTTCGGAATTCGAAATTTGAAAGTG 1686
Db 1627 CTGCATGACGAGGATGACACCAAGGCTTAACTTCGGAATTCGAAATTTGAAAGTG 1686
QY 212 AsnAsnLeuLysSerLysAsnAlaGlnGlnIleHisIleLysPheGlySerCysThrAsn 231
   1687 AATATCTAAAGCTAAAGTAAATTCGACCAACAAATTCATATCAATTTGATGATGACTTAAT 1746
Db 1687 AATATCTAAAGCTAAAGTAAATTCGACCAACAAATTCATATCAATTTGATGATGACTTAAT 1746
QY 232 ValValAlaSerAsnLeuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyVal 251
   1747 GTTGTAGCTTCAAAATTTGATGATCAATGCTTCAGCAAGGCCCAAAATTCGATGAGATC 1806
Db 1747 GTTGTAGCTTCAAAATTTGATGATCAATGCTTCAGCAAGGCCCAAAATTCGATGAGATC 1806
QY 252 HisValSerAsnThrGlnTrrIleGlnIleSerAspTrrIleIleGlyThrGlyAspAsp 271
   1807 CATGTATCAAAATTCATCAATATATTCGAAATTCGATACATATTAATGGAACAGTATGAT 1866
Db 1807 CATGTATCAAAATTCATCAATATATTCGAAATTCGATACATATTAATGGAACAGTATGAT 1866
QY 272 CysIleSerIleValSerGlyAsnValAlaGlnAlaThrAsnIleThrCysGlyPro 291
   1867 TGTATTTCAATTTGCTTCTGATCTCAAAATGTCAGGCCCAAAATTTTCTTGTGTCCA 1926
Db 1867 TGTATTTCAATTTGCTTCTGATCTCAAAATGTCAGGCCCAAAATTTTCTTGTGTCCA 1926
QY 292 GlyHisGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTrrValSerAsn 311
   1927 GGTCAATGATTAAGATTAATGGAAGCTTAGATCTGGAATTCAGAAAGCTTATGTCTAAT 1986
Db 1927 GGTCAATGATTAAGATTAATGGAAGCTTAGATCTGGAATTCAGAAAGCTTATGTCTAAT 1986
QY 312 ValThrValAsnGlnAlaLysIleIleGlyValGlnAsnGlyValArgIleLysThrTrr 331
   1987 GTTACTGTAAATGAAGCCAAATTTTCTGATGCCGAAATGAGATTAAGATTAAGCTTGG 2046
Db 1987 GTTACTGTAAATGAAGCCAAATTTTCTGATGCCGAAATGAGATTAAGATTAAGCTTGG 2046
QY 332 GlnGlyGlySerGlyGlnAlaSerAsnIleLysPheLeuAsnValGlnMetGlnAspVal 351
   2047 CAGGAGAGATCTGACCAAGCTAGCAACATCAAAATTTTCTGAATGTGAAATGCAAGACGTT 2106
Db 2047 CAGGAGAGATCTGACCAAGCTAGCAACATCAAAATTTTCTGAATGTGAAATGCAAGACGTT 2106
QY 352 LysTrrProIleIleIleAspGlnAsnTrrCysAspArgValGlnProCysIleGlnGln 371
   2107 AAGTATCCCATTAATTAATGAACCAAAATTAATGTGTGATCAAGTGAACCATTAATCAACAG 2166
Db 2107 AAGTATCCCATTAATTAATGAACCAAAATTAATGTGTGATCAAGTGAACCATTAATCAACAG 2166
QY 372 PheSerAlaValAlaGlnValLysAsnValValTrrGlnAsnIleLysGlyThrSerAlaThr 391
   2167 TTTTCAGCAGCTTCAAGTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2226
Db 2167 TTTTCAGCAGCTTCAAGTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2226
QY 392 LysValAlaIleLysPheAspCysSerThrAsnPheProCysGlyGlyIleIleMetGln 411
   2227 AAGGTGGCCATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2286
Db 2227 AAGGTGGCCATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2286
QY 412 AsnIleAsnLeuValGlyGlySerGlyLysProSerGlyAlaThrCysLysAsnValHis 431

```

```

Db 2287 AATATAAATTTAGTAGGGAAGTGAAGAAACCATCAGAGGCTACGTGCAAAATGTCCAT 2346
QY 432 PheAsnAsnAlaGlnHisValThrProHisCysThrSerLeuGlnIleSerGlnAspGln 451
   TTTTAAACATGCTGAAATCATGTATACCAACATGCTGCACTTCTAGAAATTTTCAGAGATGAA 2406
Db 2347 TTTTAAACATGCTGAAATCATGTATACCAACATGCTGCACTTCTAGAAATTTTCAGAGATGAA 2406
QY 452 AlaLeuLeuTrrAsnTrr 457
   GCTCTTTTGTATTAATTAAT 2424
Db 2407 GCTCTTTTGTATTAATTAAT 2424

```

## RESULT 2

```

US-10-691-374-1
: Sequence 1, Application US/10691374
: Publication No. US20040250322A1
: GENERAL INFORMATION:
: APPLICANT: McCallum, Claire
: APPLICANT: Slade, Ann J.
: APPLICANT: Colbert, Trent
: APPLICANT: Knaut, Vic
: APPLICANT: Anawah Inc.
: TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
: FILE REFERENCE: MHB 02-276
: CURRENT APPLICATION NUMBER: US/10/691,374
: NUMBER OF SEQ ID NOS: 50
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 7456
: TYPE: DNA
: ORGANISM: Lycopersicon esculentum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1479)..(1757)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (3327)..(3491)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (3696)..(3716)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (4260)..(4467)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (4567)..(4648)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (5602)..(5710)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (6139)..(6255)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (6788)..(7045)
: OTHER INFORMATION:
US-10-691-374-1

```

## Alignment Scores:

Pred. No.:	2.48e-163	Length:	7456
Score:	1506.00	Matches:	454

Percent Similarity: 24.43% Conservative: 0  
Best Local Similarity: 24.43% Mismatches: 3  
Query Match: 63.01% Indels: 1404  
DB: 18 Gaps: 8  
US-10-691-374-2 (1-457) x US-10-691-374-1 (1-7456)

QY 1 MetValIleGlnArgAsnSerIleLeuLeuIleIleIlePheAlaSerSerIleSer 20  
DB 1479 ATGGTATATCAAGAAAGATAGTATTCCTTCCTCATATATATTTTGCCTCATCAATTTCA 1538  
QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheValIleValIleAsnIleLeu 40  
DB 1539 ACTGTAGACGATATTTATGTAGCAATTTATTAACAAGTTTATGATATATTTCTT 1598  
QY 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerIleAsnIleGlu 60  
DB 1599 GAACAAGATTTTGCCTCATATTTTCAAGCTTATCTTTATTTGACAAAATATTTGA 1658  
QY 61 SerAspAsnAsnIleAspIleValIleAspIleValIleValIleAsnValIleuSer 80  
DB 1659 AGCAACAAATAAATATGACAAAGTTGATMAAAATGGGATTAAGTATATGATCTTAC 1718  
QY 81 PheGlyAlaIleGlyValAspGlyIleThrTyrAspAsnIle----- 93  
DB 1719 TTGGAGCTAAGGGTATGAGAAACATATGATATATTTGTAATTTAAATATTGAA 1778  
QY 93 ----- 93  
DB 1779 TATATTTGTGGGATGAAAATGATAGAAATATTAAGATTAATTTGGAAAGATGAAAAGTT 1838  
QY 93 ----- 93  
DB 1839 ATATTTTATAAAGTAAATAATATTTCTCGTTTTAGTATTAAGGTGAAATGAGTTTC 1898  
QY 93 ----- 93  
DB 1899 TCGTTAAGCGAGAAAAGCTATTTTCCATGCTACTGATTTTTTTTTTACTTTTAATA 1958  
QY 93 ----- 93  
DB 1959 CGTCATAGTATTTGCTATCTCAAGATTAAGACACTTATTAATGATGATTTAGTCTGA 2018  
QY 93 ----- 93  
DB 2019 AAAGAAATGATAGTAAATTTTGTCTAATATTAATCAATTTCTTATAGTATATTTTC 2078  
QY 93 ----- 93  
DB 2079 AACCAAAATACAAAGCTAATCCATTAAGTGGGCTCTAGAAATAAAGATAGTTCTAT 2138  
QY 93 ----- 93  
DB 2139 TCAATTTCTAACCTTATTTAATTTTAGTGGAAACCTGCAAAAGCAAAAGTATTC 2198  
QY 93 ----- 93  
DB 2199 AAACCTTTATATTCGGAATTCGAGCAACAATATGACAACTCACAATGATATAGT 2258  
QY 93 ----- 93  
DB 2259 CCTAATATATTAATTTTCTAATAAAATATCTTCAATCTACATTAATGAAATTTGAAA 2318  
QY 93 ----- 93  
DB 2319 ATGACTTTTATCTCTATCGAACACATATCAAGATTTCTTTAAGAATTTACACATACAT 2378  
QY 94 -----AlaPheGluGlnAlaIlePheGlu 101  
DB 2379 TTGGATATGTTTCTTATCGTGAATAATATCTTTCAAGCAATTTGACAGCAAGTATATA 2438  
QY 101 uAlaCysSerSerArgThrProValGlnPheValIleProValAsnIleValAsnTyrLeuLeu 121

DB 2439 AGCATGTTCACTAGAACACCTGTTCAATTTTGTTGTTCTCTAATAAACAAAGATTAATCTTCT 2498  
QY 121 uIleGlnIleThrPheSerGlyProCysArgSerSerIleSerValIle----- 137  
DB 2499 CAAGCAAAATCACCTTTTCAGTCTCAATGACATCTTCTATTTCCAGT-AAAGTTAGCATAT 2557  
QY 137 ----- 137  
DB 2558 TCATATTTATATCTCTTTGTAGCAATATATATCTGTTTAGCAAAATTTAAGAA 2617  
QY 137 ----- 137  
DB 2618 AGTAAATCAAGATAGATAAACAATGAATTTTGTCTACTAATTTTAGCGGATTTAGTGAGAA 2677  
QY 137 ----- 137  
DB 2678 TTATCAAAATGTATAGTATAGTATGACAACTTAGCTATGAATTAGCTAGAGAAAGTT 2737  
QY 137 ----- 137  
DB 2738 TGAATCAATCTATATTTTTTTTTTGTAGATTAAGATATTTGAACACATGATTAATTA 2797  
QY 137 ----- 137  
DB 2798 TTAATTAATCTTAATTAATATATGTCATAGATAGTTCAAACTAAAGACTGCATAAGAA 2857  
QY 137 ----- 137  
DB 2858 AATAAGAAAGAAATATTTATTTTAAATAATTAATAAAGAAATATGAGAAATTAATTT 2917  
QY 137 ----- 137  
DB 2918 CAAAGCGAAGATATACATATCTTAGGGGATTAAGATTAATATATATGTAAGAAA 2977  
QY 137 ----- 137  
DB 2978 CAGCACTACATATCTAATAAAGTCTAATAAGTATTAATAAATAATAGTGTATAGCAA 3037  
QY 137 ----- 137  
DB 3038 CAGTTATCCCTACAAAACCTTTTGTGGGTAGATCGATCCAGAGTTGTTCCAGACTCT 3097  
QY 137 ----- 137  
DB 3098 TCCTTAATAAAATGTTTTTCTAATAATAATTGAAGAAATGTTATATGATGAATAATAT 3157  
QY 137 ----- 137  
DB 3158 GAAAGAAAACATATCAATTTAATAAATAATAAAGTAAATCAAGTAACGAATTAACAAATA 3217  
QY 137 ----- 137  
DB 3218 GGAATATATCTATTAATGAATAATTTAGTGGCTTTGTTAATCATATCTTAGTTATTC 3277  
QY 138 -----IlePheGlyIle 141  
DB 3278 ATGTGTTCTTAATTTCCCTCTTATTTTTTTGAATATACTAATGACAGATTTTGGAGTC 3337  
QY 141 rIleuGlnIleSerSerIleSerAspTyrIleAspArgIleIleIleIleIleIleIleIleIle 161  
DB 3338 CTTAGAGCATCTAGTAATAATTTCAAGCTACAAAGATTAAGAGCTTTGGATTTGCTTTTGA 3397  
QY 161 pSerValGlnAsnLeuValIleGlyIleGlyIleThrIleAsnGlyAsnGlyGlnValIle 181  
DB 3398 TAGTGTTCAAAATTTAGTGTGGAGAGAGAACTATCAATGCGAATGACAAAGTATAG 3457  
QY 181 pThrProSerSerCysIleIleAsnIleSerIle----- 192  
DB 3458 GTGGCCAAAGTTCTTGCAAAATTAATAATCACTGTAATTTTAATACTTGTCTTATAGT 3517  
QY 192 ----- 192  
DB 3518 TTACGCTATGTTGCTCGAATTTCTTAACCTTGTCTAAGATATTAATATTTTGAAGAA 3577

```
OY 192 ----- 192
Db 3578 GGTGTCACAAATGCATCATATTTTGTAGAGATTCGACCAATATATGTTATGTATATCTA 3637
OY 193 ----- Pr 193
Db 3638 ATTTTCAGAGCATCTTTCCTGTGATCTGATGATTTTACCTTTTTCCTTCATGACGCC 3697
OY 193 oCyArgspAlaProThr ----- 199
Db 3698 ATGCGAGGATCCACCAACGATGACGTATATTCATTTGATTAATAAAAAAGCCTA 3757
OY 199 ----- 199
Db 3758 AAATATATTTGAATTTTAATTGAAAAGTTATTAATTTCTTAACCTTTGGCGAGACCTAT 3817
OY 199 ----- 199
Db 3818 TACCCCTTGACATTTTAATAGTGTATTTTAAAGATATAAAGTTAGTTGAACAA 3877
OY 199 ----- 199
Db 3878 AATTAGATATTCAAAACTATTTGAAAATTAATAAATGCAATTTTTCATATCA 3937
OY 199 ----- 199
Db 3938 ATATGATTAATAAATAATTAGTTAAAGTTCTTATGATTTGATTCATAAAAAATCAATCANG 3997
OY 199 ----- 199
Db 3998 ACAAAACATAGTAGACGAGAAAGTATATACATACCTTCAGATGACATGATTTGT 4057
OY 199 ----- 199
Db 4058 ACAACACCTCAAAACCTAGCTTTTCTTGATTTATATTTCTTATTTCTTTAATATGTA 4117
OY 199 ----- 199
Db 4118 TCAAAGCTATTTAGTTCTGTCAAAATCTATACATGGAACCTATCTTGAAGCTGCT 4177
OY 199 ----- 199
Db 4178 ACATTCGAGATCGTTGAACATGATGATTAATTTAATTTGATTTAATATTA 4237
OY 200 ----- AlaLeuThrPheTyrPheAsnLysValAs 212
Db 4238 ACTAATATTTGTTAATTTTCAGGCTTACCTTCTGATTTCCAAAAATTTGAAAGTGA 4297
OY 212 nasLeuLysSerLysAsnAlaGlnGlnIleHisLysPheGluSerCysThrAsnVa 232
Db 4298 TAACTCAAAAGAGTAAATAATGCAACAAATTCATATCAAAATTTGAGTCATGACATAATGT 4357
OY 232 lValAlaSerAsnLeuMetIleAsnAlaSerAlaLysSerProAsnThrAsnGlyValHis 252
Db 4358 TGTAGCTTCAAAATTTGATGATATCAATGCTTACCAAAAGAGCCCAAAATCTGATGAGTCCA 4417
OY 252 sValSerAsnThrGlnTyrIleGlnIleSerAspThrIleIleGlyThr ----- 268
Db 4418 TGTATCAAAATCTCAATATATTCAAATATCTGATCTATTAATGGAACAAGTTATTTAT 4477
OY 268 ----- 268
Db 4478 TTAATTTTATTTATTCATTAATTAAGAAAAAGAGATTTTATTTGATTAATCTA 4537
OY 269 ----- GlysAspSerCysIleSerIleValLysGlyse 279
Db 4538 AATATATATTTTATTTTATTTTATATAGATGATGATTTGATTTCAATTTGTTCTGATC 4597
OY 279 rGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSer ----- 296
Db 4598 TCAAAATGTGACGGCCAAATATATTAATCTTGTGTCCAGGTGATGTATTAAG-GTACTCTA 4656

OY 296 ----- 296
Db 4657 TTTTACAAATATACCTTTCCATTTTCTCTATTTTCATATAAAGATATGATATATATA 4716
OY 296 ----- 296
Db 4717 TTACTTTAAATCCTTAATTAATTAATTAATGCAAAATTTTTCCTTGTCTTATGTTAAT 4776
OY 296 ----- 296
Db 4777 GACTTAGCACAATAATTAAGGCCGTTTGATGGCGGAATTAAGACGCTTTAAAAAGTA 4836
OY 296 ----- 296
Db 4837 CTTTAAAAAGTGTGAACTTATTTTAAATAATAGCAGTTATTCGGTTTGATTAAGTGC 4896
OY 296 ----- 296
Db 4897 TGAAGTTGTATGTCAAAAGTGAAAAAGGAAAAATGGAAGAAAGATGTAGGTTATA 4956
OY 296 ----- 296
Db 4957 TGGTATTTGTATATAAATAATTAAGCAGAAAAAGATAAATGTGTCACTTAACA 5016
OY 296 ----- 296
Db 5017 ACTTAAAGTACCTACCTACCTACCCAGCTTTTAATCTTTGGCTTAAATAATGTTT 5076
OY 296 ----- 296
Db 5077 TTTTAAACCTTAATAAAGTTGTTTGATTTGCCAAGAGCTTAATATGCAAAAACC 5136
OY 296 ----- 296
Db 5137 AGCTTTAAGTCACTTTGACAGCTTTTAAGCTGAGCAACAGAGCTTTAAATGTCTG 5196
OY 296 ----- 296
Db 5197 CTTAGATGTGTATATATTAATTTAGCTTTTGTGAAGTATATATCTTAAGTTCAA 5256
OY 296 ----- 296
Db 5257 CATAAATATACATGCTTTAATACATGACATATAGTTAATCAAAAAGCAATGATGAATA 5316
OY 296 ----- 296
Db 5317 TTTTGCAAATTTGATTAATTCACAAAGAAAGGATAGTTCAAGTATCAATTCATGAAT 5376
OY 296 ----- 296
Db 5377 TGAAGATATCATAAAGCTAAATTAAGAAATCAATTAATGAGGATCAAAAATGTTAT 5436
OY 296 ----- 296
Db 5437 TACCTATTAATAATACATATTCATTTTCATATTAATTAATTAAGAGTTTTATA 5496
OY 296 ----- 296
Db 5497 ATCTAATTAACATGCAATTAATTAATGAGAAATGTGTTTGATCATATATCTTCT 5556
OY 297 ----- IleGlySerLeuG1 301
Db 5557 GAATATTTGCTATTTTCTCTTTTATTTTTCATGAGATTAATGATTTGAAAGCTTAG 5616
OY 301 ySerGlyAsnSerGlyAlaTyrValSerAsnValThrValAsnGlyAlaLysIleIleG1 321
Db 5617 ATCTGAAAATTCAGAGCTTATGTGTCTAATGTTACTGTAATTAAGAACCAAAATATATCG 5676
OY 321 yAlaGluAsnGlyValArgIleLysThrThrGln ----- 332
Db 5677 TGGCAAAATGAGATTAGATCAAGACTTGGA-GATACCTTCCCCCCCCCCCCCCCC 5735
OY 332 ----- 332
```



```

Db      5736 CACAGGCCATTTTAAATTTTAAATTTTATGGAATATCATATTAAGATTA 5795
Qy      332 -----
Db      5796 ATTGATTTCAATGTTGAATTTATATTTGATAAGATGATATTTTACTAGCTTCTAT 5855
Qy      332 -----
Db      5856 GTTATATAGAAAAAAATGTTGAGAACTTCAGATTATTTGTACTGTAAGTAAAT 5915
Qy      332 -----
Db      5916 GTTGCTTTGTTAGAAAGTTGGTTATCCAGTTTGGGTCATGATTAACCAACTTA 5975
Qy      332 -----
Db      5976 TAATGAAAAGGGCTGCAACGGCCGCCACTAGTCTAGTATCAATAGGAAGATCTAC 6035
Qy      332 -----
Db      6036 GTCTGTTATTCAGATGACGTTCTTGTTGAATGTTAATTAATTAATTAAACA 6095
Qy      333 -----
Db      6096 TGTATTAAGCATATATTAATTAATGCTTTAATATGATGAGGAGATCTGACACAC 6155
Qy      338 aSer[en]1elysPheLeuAsnValGlnMetGlnAspValLysTyrProIleIleAs 358
Db      6156 TAGCAACATCAATTTCTGAAATGTGGAAATGCAAGCGTTAAGTATCCATATTTATGA 6215
Qy      358 pGln[en]TyrCysAspArgValGlnProCysIleGln-----GlyGlySerGlyGlnAl 338
Db      6216 CCAAAACATATGATGATGAGTGAACCATGATATACAAAGGTAATTTTATTAACGAAC 6275
Qy      370 -----
Db      6276 AATTTATATATTTTATTACTTCTTAATCACCCTTACATTAATAACTTGAATCTT 6335
Qy      370 -----
Db      6336 TTCACAGTAGTAACCTTTTGAATGATTTTATGTAATGATTCATTATTCCTTTTA 6395
Qy      370 -----
Db      6396 TTTTCTTCTAATTAATGATCTTTTGAATGATGCTTAATAAATCTGTTAAAGTAACT 6455
Qy      370 -----
Db      6456 GAATATCATAGAAAAAATGTTAGATTATATCTAAATTTTATTAATTAAGACGTT 6515
Qy      370 -----
Db      6516 ATCTAATATTTTGTATGTAAGATGGAACAATATACATTAATAACATTGATTCAAATTTA 6575
Qy      370 -----
Db      6576 ATAAATATCTAAATAATGATTCAAATCAATCATGACTACAAACGAATACAGATG 6635
Qy      370 -----
Db      6636 CTCAAACATATAGATGAAGTCAATTCAAAACGAATCAAAATATAGTATATCTTTAA 6695
Qy      370 -----
Db      6696 AAGAGGCAATTTGGTAAATAAGTAAATACTTAAGTATTAATAAATCTCAACTCGA 6755
Qy      371 -----
Db      6756 TCTCTCAGCATTAATTAATCACTTTGTTCCAGTTTTCACAGTTTCAGGAAAGTGG 6815
Qy      381 aIlyrGlnAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAspCysSerT 401

```

```

Db      6816 TGTATGAGATATCAAGGGCACAAGTCACAAAGGTGGCCATTAATTTGATTCGACGA 6875
Qy      401 hAsnPheProCysGlyGlyIleIleMetGlnAsnIleAsnLeuValGlyGlySerGlyL 421
Db      6876 CAACCTTCCATGATGAGGAATTAATGAGAAATATTAATTTAGAGGGAAGTGA 6935
Qy      421 yProSerGlnAlaThrCysLysAsnValHisPheAsnAsnAlaGluHisValThrProH 441
Db      6936 AACCATCAGAGGCTACGTGCACAAATGTTCCATTTTAAACAATGCTGAACATGTTACACAC 6995
Qy      441 iCysThrSerLeuGluIleSerGluAspGluAlaLeuLeuTyrAsnTyr 457
Db      6996 ACTGCACCTTCACTAGAAATTTTCAGAGGATGAAGCTCTTTTGTATATATAT 7045

RESULT 3
US-10-691-374-47
; Sequence 47, Application US/10691374
; Publication No. US20040250322A1
; GENERAL INFORMATION:
; APPLICANT: McCallum, Claire
; APPLICANT: Slade, Ann J.
; APPLICANT: Colbert, Trent
; APPLICANT: Knauft, Vic
; APPLICANT: Anawah Inc.
; TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
; FILE REFERENCE: MBH 02-276
; CURRENT APPLICATION NUMBER: US/10/691,374
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 7456
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1479)..(1757)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2416)..(2547)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3327)..(3491)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3696)..(3716)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4260)..(4467)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4567)..(4648)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5602)..(5710)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6139)..(6255)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6788)..(7045)
; OTHER INFORMATION:
US-10-691-374-47

```

## Alignment Scores:

Pred. No.:	2,12e-162	Length:	7456
Score:	1498.00	Matches:	453
Percent Similarity:	24.38%	Conservative:	0
Best Local Similarity:	24.38%	Mismatches:	4
Query Match:	62.68%	Indels:	1404
DB:	18	Gaps:	8

US-10-691-374-2 (1-457) x US-10-691-374-47 (1-7456)

```

QY      1 MetValIleGlnArgAsnSerIleLeuLeuIleIlePheAlaSerSerIleSer 20
Db      1479 ATGGTATCCAAAGGATAGTATCTCCTTCATATATTTTGGCTTCATTCATTTCA 1538
QY      21 ThrCysArgSerAsnValIleAspAspAsnLeuPheIleValIleAspAsnIleLeu 40
Db      1539 ACTTGTAGAGCAAGATTTATGATGACATTTATTCACAAAGTTATGATTAATATCTT 1598
QY      41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerTyrAsnIleGlu 60
Db      1599 GAACAAAGATTTGCTCATGATTTTCAGCTTATCTTCTTTTGGAGCAAAATATTGAA 1658
QY      61 SerAsnAsnAsnIleAspIleValAspIleValIleValValIleAsnValIleuSer 80
Db      1659 AGCAACATTAATATTTGACAGGTTGATTAATAATGGATTAAGGATTAATGACTTAC 1718
QY      81 PheGlyAlaIleValAspGlyIleValThrTyrAspAsnIle----- 93
Db      1719 TTTGGAGCTAAGGGTGAATGAAACATATGATTAATTTGTAAGTATTAAATATTGAA 1778
QY      93 ----- 93
Db      1779 TATATTTGGGGATGAAATGATAGAAATATAGAAATATTGGAAGATGAAAGTT 1838
QY      93 ----- 93
Db      1839 ATATTTATAAGTAAAGTAAATATTTTCTGCTTTTATAGTAAAGTAAAGTTC 1898
QY      93 ----- 93
Db      1899 TCGTTAAGCGAGGAAAGCTATTTCCATGTTAAGTATTTTCTTACTTTAATA 1958
QY      93 ----- 93
Db      1959 CGTCATAGTATTTGCTATCTACTCAAGATTAAGACATTAATTAATGATTAAGTCTCGA 2018
QY      93 ----- 93
Db      2019 AAGAAATGATAGTATTTTGTCTAATTAATCTATCATTTCTTAATATATTTTTC 2078
QY      93 ----- 93
Db      2079 AACCAAAATTAACAAGCGTAAATCAATTAAGTGGGCTTAAGATAAGTAAAGTTCTAT 2138
QY      93 ----- 93
Db      2139 TCAATTTCTAACCTTATTTAATTTTAGTGAACCTCGACAAAACGAACAAAGTATTC 2198
QY      93 ----- 93
Db      2199 AAACTTTATATTCGAATTCGAGACCAACATATGAACAACTCACACATGCAATATAGT 2258
QY      93 ----- 93
Db      2259 CCTAATATATATATTTTCTTAATAAAATATCTTCATCTACCATATTGAATATTGAAA 2318
QY      93 ----- 93
Db      2319 ATGACTTTATTCCTATCGACACATATCAAGAGTTCTTTTAAGATTAACCATCAT 2378
QY      94 ----- 1AlaPheGluGlnAlaTyrPheGly 101
Db      2379 TTGGATGTTCTTATGCTGTTAAATATATCTTTCAGGCAATTGAGCAAGCATGGAATGA 2438

```

```

QY      101 uAlaCysSerSerArgThrProValGlnPheValValProIleAsnIleValIleLeu 121
Db      2439 AGCATGTTCACTAGAACACCTGTTCAATTTGTTGTTCTTAATAACCAAGATTTATCTTCT 2498
QY      121 uLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValys----- 137
Db      2499 CAAGCAATACACCTTTTCAAGTCCATGCAAGATCTTATTTCAAGT-AAAGTTAGCAATAT 2557
QY      137 ----- 137
Db      2558 TGATTTATATATCTCTTTGTAGCAATAATATATCTGTTATAGACAAAATTAAGA 2617
QY      137 ----- 137
Db      2618 AGTAATCAAGAATAGATAAACAATGAATTTTCGTCACTAATTTAGCGGATTAGTAGAGAA 2677
QY      137 ----- 137
Db      2678 TATCAAAATGTATGTAGTATGACATGACACTTAGATTAATTAAGTATTAAGTAAAGTT 2737
QY      137 ----- 137
Db      2738 TGATGCTAATTTCTATTTTGTGTAGATAAGATAATTTGAACAACATGATTAATTA 2797
QY      137 ----- 137
Db      2798 TTAATTAATCTTAATTAATATGTCATGATGACTTCAAACTAAGACTGTCAAAAGAA 2857
QY      137 ----- 137
Db      2858 AATAAGAAAGAAATATTATTTTAAATAATTAATAAGAAATAATGAGAAATAAATT 2917
QY      137 ----- 137
Db      2918 CAAGCGAAGATTAATCAATATCTATGGGATTAAGATTTATATATGTTAAGAAA 2977
QY      137 ----- 137
Db      2978 CAGACTACACATATCTAATTAAGTCTCATTAATGATTAATAAATAATGATGTAAGCA 3037
QY      137 ----- 137
Db      3038 CAGTTATCCCTAAGAAAACCTTTTGGGGTATGATGATCAGAGTGTTCACAGCTCT 3097
QY      137 ----- 137
Db      3098 TGCTTAATAAAATGTTTTCTAATTAAGTTGAAGAATGTTATATATGAGAAAATAT 3157
QY      137 ----- 137
Db      3158 GAAGAAAACATATCAATTTTAATAAATAATTAAGTAATCAAGTAACGAATAAACAATA 3217
QY      137 ----- 137
Db      3218 GGAATATACATATAATGAATAATTAAGGCTTTTCGTTAACAATATCTAGTTATTC 3277
QY      138 ----- 1IlePheGlySe 141
Db      3278 ATTTGTTCTTAATTTCCCTCTTAATTTTGTGAATTAATGACAGATTTTGGATC 3337
QY      141 rLeuGlnAlaSerSerIleSerAspTyrIleAspArgArgLeuTrrIleAlaPheAs 161
Db      3338 CTTAGAGCATCTAATAAATTTTCAGATCAAAAGATGAAAGGCTTTGGATTTGCTTTGA 3397
QY      161 pSerValGlnAsnLeuValValGlyGlyGlyThrIleAsnGlyAsnGlyGlnValTr 181
Db      3398 TAGTGTTCAAAATTTAGTTGTGGAGGAGGAATCATCATGCGCAATAGACAAGATAG 3457
QY      181 pTyrProSerSerCysIleAsnIleAsnIleSerLeu----- 192
Db      3458 GTGGCCAAGTTCTTGCAAAATTAATATCACTGTATTAATTTATTAACCTTGCTTAATAGT 3517

```

QY	192	-----	192	-----	192
Db	3518	TTTACGCTATGTTGCTCGAATTCCTTAACTGTCTTAAGATATATATATTTGAAGA	3577		
QY	192	-----	192	-----	192
Db	3578	GGTGTCAAAATGCATCACTATTTAGAGATTCCGACCAATATTAGTTTATGTATCTA	3637		
QY	193	-----	193	-----	193
Db	3638	ATTTTCAGAGCATCTTGGCTGTACTGATCATCTTTACCTTTTCTTCATGCAGACC	3697		
QY	193	OCy <sup>h</sup> ArgAp <sup>h</sup> AlaPro <sup>h</sup> Thr-----	199		
Db	3698	ATGCAGGATGCACCAACGGTACGTAAATGCAATTTGATATAAAAAAAGCCTA	3757		
QY	199	-----	199	-----	199
Db	3758	AAATATATTGAATTTAATTGAAGATTATATATCTTATCTTGGCAGAGCCTAT	3817		
QY	199	-----	199	-----	199
Db	3818	TACCCCTTGCACTATTTAATAGTATTTTAAAGATATATAAGTGTAGTTGAACAAA	3877		
QY	199	-----	199	-----	199
Db	3878	AATTTAGATATTCAAAACATATTTGAAATTACTAATTAATGCAATTTTGTGCATATCA	3937		
QY	199	-----	199	-----	199
Db	3938	ATATGATTAAAAATATATAGTTAAAGTTCTTATGATTTGATTTCAAAATATAATCATG	3997		
QY	199	-----	199	-----	199
Db	3998	ACAAACATATAGTACGAGAAAGATATATACAAATCCTCTTCAAGTGAATGCATTTGT	4057		
QY	199	-----	199	-----	199
Db	4058	ACACACACCTCAAAACCTACGTTTCTTCGATTATATTTCCATATTCCTTAAATAGTAA	4117		
QY	199	-----	199	-----	199
Db	4118	TCAAAGCTATATAGTTCTGTCAAAATCTATACATTGAAACCTATCTTGAAGCTGTG	4177		
QY	199	-----	199	-----	199
Db	4178	ACATTCGAGATCGTTGAACATGCATGAATGATTTTATACCTTGTATTTAATATTTAA	4237		
QY	200	-----	212	-----	212
Db	4238	ACTATATATTGTTAATTTTTCAGGCTTTAACCTTTCGGAATTCAGAAAAATTTGAAGTGA	4297		
QY	212	nAsnLeuLySerLy <sup>h</sup> AsnAlaGlnGlnIleHisIleLy <sup>h</sup> PheGluSerCy <sup>h</sup> ThrAsnVa	232		
Db	4298	TAATCTAAAGGTAAATAATGCACCAAAATTCATATCAAAATTTGGTATGCATGCATTAAGT	4357		
QY	232	IValAlaSerAsnLeuMetIleAsnAlaSerAlaLySerProAsnThrAspGlyValHis	252		
Db	4358	TGTACCTTCMAATTTGATGATCATAGCTTCACGAAAGGCCCAATATCTGATGGATCCA	4417		
QY	252	eValSerAsnThrGlnTy <sup>h</sup> IleGlnIleSerAspThrIleIleGlyThr-----	268		
Db	4418	TGTATCAATATCTCAATATATTTCAAAATATCTGATCTATTTATGGAAACAGGTTATTTAT	4477		
QY	268	-----	268	-----	268
Db	4478	TTAATTTTATTTATCCAAATTAATTAAGAAAAAAGAGATTTTATTTGATATCTA	4537		
QY	269	-----	279	-----	279
Db	4538	AATTAATTAATTTTATTTTATTTTATATAGTGATATGTATTTCAATTTGTTTCTGATC	4597		
QY	279	rgInaenValGlnAlaThrAsnIleThrCy <sup>h</sup> GlyProGly <sup>h</sup> IleGlyIleSer-----	296		

Db	4598	TCAAAATGTGAGGCCAACAAATATTACTGTGTCCAGGTGATGTATTAAG-GTACTCTTA	4656
QY	296	-----	296
Db	4657	TTTTCACAAATATACTGTGTTCCATTTTCTCATTTTCATTAAGAAGGTAGATGATATATTA	4716
QY	296	-----	296
Db	4717	TTACTTTAAATCCTTAATTAATTAATTGCAAAATTTTTTCTCTTGCTTATGTGTAAAT	4776
QY	296	-----	296
Db	4777	GACTTAGACAATAATTAAGGCCGTTGGATGGCGCATTAAGACACTTTAAAAAGTA	4836
QY	296	-----	296
Db	4837	CTTTTAAAAAGTGTGAACCTTATTTTAAAAATAGCAGTTATCGGTTTGGATAAAGTGC	4896
QY	296	-----	296
Db	4897	TGAAGTGTATATGTCAACGTGAAGGAAAAATGAGAAAGAAATGTTAGGTTATA	4956
QY	296	-----	296
Db	4957	TGGGTATTGTATAAAAATTAATAGCAGAAAAAGATAAATGTGTCACTTAACA	5016
QY	296	-----	296
Db	5017	ACTTATTAAGTACCTTACCCTACCCAGCTTTTAACCTTTGGCTTAAATAAGTTTTTT	5076
QY	296	-----	296
Db	5077	TTTTAAAACTTAAAAATAAGTTGTTTGTAGTATGCAAGAGCTAAATATGCAAAAAC	5136
QY	296	-----	296
Db	5137	AGCTTTTAAGTCAGTTTGACCGAGCTTTTAAGCTGAGCCAAACAGGCTTTAAAAATGCTG	5196
QY	296	-----	296
Db	5197	CTTAGAGTGTCTAATATATATTGAGCTTTTTTGAAGTATATTAATCTTAAAGTTCA	5256
QY	296	-----	296
Db	5257	CATAAATATACATGCTTTAATACATAGACATATAGTTAATCAAAAGCGAAATGATGAATTA	5316
QY	296	-----	296
Db	5317	TTTTGCGAATTTGATTAATTCACAGAAAAAGGATAGTTCAAGTGTACATTTCAATGAAAT	5376
QY	296	-----	296
Db	5377	TGAAGATATCATTAAGACTAAATTAAGAGATCAATATATGAGGATCAAAAAATGTTAT	5436
QY	296	-----	296
Db	5437	TACCTTAATTAATATCTATTCATTTTCATATTAATTAATTAAGATGTTTTATA	5496
QY	296	-----	296
Db	5497	ATCTAATTAACATGCAATATTAATTAATGCGAAATGTGTTTGGTACCTATATCTTCT	5556
QY	297	-----11eg1ysert1eug1	301
Db	5557	GAAATATTTGCTCTATTTTTTCTCTTTTATTTTCCATGATTAATCTAATAGAGCTTAGG	5616
QY	301	yserc1yAsnserglunlatyrValserAsnValThrValAsnslunlatysle1leg1	321
Db	5617	ATCTGGAATTCAGAGCTTATGTCTTAATGTTACTGTAAATGAGCGAAATTAATTCGG	5676
QY	321	yAlaGluAsnsllyValArg11elysrThrTTPGln-----	332

Db 5677 TGCCGAAATAGAGTTAGATCAAGACTTGGA-GGTACCTCCCCCCCCCCCCCCCC 5735  
QY 332 ----- 332  
Db 5736 CACAGGCCCATTTTTTAAATTTTTTAAATTTTATTCGAATATCAATATTAAAGATT 5795  
QY 332 ----- 332  
Db 5796 AATTGATTCATGTTTGAATTTATATTGATTAAGATGATGATTTTACTAGCTTCTAT 5855  
QY 332 ----- 332  
Db 5856 GTTATATAGAAAAAAATGTTGAGAACTTCAGATTATTGTAAGTCTAAGTAAAT 5915  
QY 332 ----- 332  
Db 5916 GTGTTGCTTTGTTTGAAGTTTGTTATTCAGTTTGGTCATGATTAAACCAACTTA 5975  
QY 332 ----- 332  
Db 5976 TAATGAAAAAGGGGTGCAACGGCGGCCACTAGTGTATGATCAATAGAGATCTCAC 6035  
QY 332 ----- 332  
Db 6036 GTCTGTTTATTCAGATGACGCTCTGTGATGATTAATTAATTAATTAATTAACA 6095  
QY 333 ----- 338  
Db 6096 TGTATTAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAACA 6155  
QY 338 ----- 358  
Db 6156 TAGCAACATCAAAATTTCTGAAATGGAAGAAAGGTAAGTATCCCATATTATTAAG 6215  
QY 358 ----- 370  
Db 6216 CCAAACTATTTGATGATGAGTTGACATGATATACACAGTATTTTATTAACGAAC 6275  
QY 370 ----- 370  
Db 6276 AATTATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6335  
QY 370 ----- 370  
Db 6336 TTCACTAGTTAGTAACCTTTTGAATGATTTTATTAATTAATTAATTAATTAATTA 6395  
QY 370 ----- 370  
Db 6396 TTTTCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6455  
QY 370 ----- 370  
Db 6456 GAATATCATAGAAAAAATGTTAGATTAATTAATTAATTAATTAATTAATTAATTAAGCT 6515  
QY 370 ----- 370  
Db 6516 ATCTAATATTTTGTATGTAAGTATGAAAGCATATACATAAAACATTAGATTCAAATTTA 6575  
QY 370 ----- 370  
Db 6576 ATATATATCAAAATATGATTAATCAATCATGACTACACAAAGAAATACATGACAGATT 6635  
QY 370 ----- 370  
Db 6636 CTCAAAATATAGATGAGTCAATTTCAAAACGAATCAAAATATAGATATATCTTAA 6695  
QY 370 ----- 370  
Db 6696 AAGAGACATTTGGTAAATTAAGTAAAAATCATTAAGTATTAATAAAATTTCTAAGTCA 6755  
QY 371 ----- 381  
Db 6756 TCTCTCAGATTATTTAATCACTTTGTTCCAGTTTCAAGATTCAGTAAATATGTGG 6815

QY 381 aITyrgluenillelygIyThrSerAlaThrIlyValAlaIlelyspheAepCyssert 401  
Db 6816 TGTATAGAAATATCAAGGGCACAAAGTCACAAAGGTGCCATTAAATTTGATTGCAGCA 6875  
QY 401 hrAnpProCyGgIyIleIleMerGluAsnIleAsnIleValGlyGlySerGlyL 421  
Db 6876 CAACCTTCATGCTGAGAGAAATTAATTAATGAGAAATTAATTAATTAATTAATTAATTAATTA 6935  
QY 421 yepProSerGluAlaThrCylysaenValHispheAsnAsnAlaGluHisValThrProH 441  
Db 6936 AACCACTCAAGGCTACGTGCAAAATGTCATTTTAACAATGCTGCAACATGTTACACAC 6995  
QY 441 iCyThrSerleuGluIleSerGluAspGluAlaLeuLeuTyraSerTy 457  
Db 6996 ACTGCACCTTCATTAAGAAATTTTCAGAGATGAAGCTCTTTGTATTAATTAAT 7045

RESULT 4  
US-10-691-374-49  
Sequence 49, Application US/10691374  
Publication No. US20040250322A1  
GENERAL INFORMATION:  
APPLICANT: McCallum, Claire  
APPLICANT: Slade, Ann J.  
APPLICANT: Colbert, Trent  
APPLICANT: Knaut, Vic  
APPLICANT: Anawah Inc.  
TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-  
FILE REFERENCE: MHBH 02-276  
CURRENT APPLICATION NUMBER: US/10/691.374  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 49  
LENGTH: 7456  
TYPE: DNA  
ORGANISM: Lycopersicon esculentum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1479)..(1757)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2416)..(2547)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3327)..(3491)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3696)..(3716)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (4260)..(4467)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (4567)..(4648)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (5602)..(5710)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (6139)..(6255)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (6788)..(7045)

## OTHER INFORMATION:

US-10-691-374-49

## Alignment Scores:

Pred. No.:	2,126-162	Length:	7456
Score:	1498.00	Matches:	453
Percent Similarity:	24.38%	Conservative:	0
Best Local Similarity:	24.38%	Mismatches:	4
Query Match:	62.68%	Indels:	1404
DB:	18	Gaps:	8

US-10-691-374-2 (1-457) x US-10-691-374-49 (1-7456)

```
QY      1 MetValIIIGlnArgAsnSerIIleuLeuLeuIIeIIlePheAlaSerSerIIleuSer 20
DB      1479 ATGGTTATCCAAAGAAATAGTATTCCTTCATATTAATTTTGGCTTCATCAATTTCA 1538

QY      21 ThrCysArgSerAspValIIleAspAspAsnLeuPheValIleValIleValIleVal 40
DB      1539 ACTTGTAGAGCAATGTTATTTGATGACATTTATTCAAACAAAGTTATGATATATTTCTT 1598

QY      41 GlnGlnIlePheAlaHisAspPheGlnAlaIleuSerIleuSerIleuSerIleuSer 60
DB      1599 GACAAAGAAATTTGCTCATGATTTTCAAGCTTATCTTTCTTATTTGAGCAAAATATTGAA 1658

QY      61 SerAspAspAsnIIleAspIleValIleAspIleAspValIleValIleValIleuSer 80
DB      1659 AGCAACAATATATTTGACAAAGGTTGATATAAAATGGATTTAAAGATATATGACTTAC 1718

QY      81 PheGlyAlaIleValIleAspGlyIleAspGlyIleAspGlyIleAspGlyIleAspGly 93
DB      1719 TTTGGAGCTAAGGGTGAATGGAABAAACATATGATATTTGTATTTGATTTAAATTTGAA 1778

QY      93 ----- 93
DB      1779 TATATTGGGGAGTAAATATGATAGATATTAAGAAATTTTGGAAAGATGAAAGTT 1838

QY      93 ----- 93
DB      1839 ATATTTTAAAGTAAATAATTTTCTCGTTTGTATTAAGGTGAAATGAGTTTC 1898

QY      93 ----- 93
DB      1899 TCGTTAAGCGAAGAAAGCTATTTCCATGTAAGTATTTTCTTACTTTAATAA 1958

QY      93 ----- 93
DB      1959 CGTCATAGTATTTGCTATCTAAGATTAAGACACTTATTTATGATGATTTAGTCTCGA 2018

QY      93 ----- 93
DB      2019 AAAAAGATTTGATGATTTTGGCTTAATTAATCAATTTCTTATATGATATTTTTC 2078

QY      93 ----- 93
DB      2079 AATCAAAATTAACAAGCTAATCAATTAAGTGGGCTTGAATTAAGATTAAGTTCTAT 2138

QY      93 ----- 93
DB      2139 TCAATTTTAACTTATTTAATTTTATGTAAGAACTCGACAAAGAAAGAAAGATATTC 2198

QY      93 ----- 93
DB      2199 AAATTTTATATTTGGAATTTGAGACCAACATATGAAACAACTCAGCATATATAGT 2258

QY      93 ----- 93
DB      2259 CCTAATATATATATTTTCTAATAAAATATCTTCAATCTACATATTAATTAATTTGAAA 2318

QY      93 ----- 93
DB      2319 ATGACTTTTATCTATCGAACACATATCAAGATTTCTTTAAGATTTTACCACTACAT 2378
```

```
QY      94 -----AlaPheGlnGlnAlaTrpAsnGly 101
DB      2379 TTGGTATGTTTCTTATCGGTAAATAATATCTTCAGGCACTTTGAGCAAGCATGCAATCA 2438

QY      101 uAlaCysSerSerArgThrProValGlnPheValIleProIleAsnIleAsnIleuSerIleu 121
DB      2439 AGCATGTTTCATCTAGAACACCGTTCATATTTGTGTTCTTAATAAACAAAGATATATCTTCT 2498

QY      121 uIleGlnIleThrPheSerGlyProCysArgSerSerIleSerValIle----- 137
DB      2499 CAAGCAATTCACCTTTTCAGGTCCATGCAATCTTCTATTTCAAGT-AAAGTTAGCATAT 2557

QY      137 ----- 137
DB      2558 TGATTTATTTATTCCTCTTTGTGACAAATATATATCTGGTTATGACAAATTTAAGA 2617

QY      137 ----- 137
DB      2618 AGTATCAAGATAGATTAACAATGAATTTTGTCACTAATTTTACGGATTAAGTGAAGAA 2677

QY      137 ----- 137
DB      2678 TTATCAAAATGTTATGTTAGCTATGAGCAACTTACCTATGAATTTAGCTATGAGAAAGTT 2737

QY      137 ----- 137
DB      2738 TGATCTAATTTCTATTTTCTTTTGTGATGATTAAGATATTTGAAACACATGATTAATTA 2797

QY      137 ----- 137
DB      2798 TTAATTTATGTTAATTAATATGATGATGATGTTCAAACTAAAGAACTGTCAAAAGAA 2857

QY      137 ----- 137
DB      2858 AATTAAGAAATATTTATTTTAAATAATTAATAAAAGAAATATGAGAAATTAAT 2917

QY      137 ----- 137
DB      2918 CAAGCGAAGATTTTACATATATCTATGAGGATTAAGAGATTTATATATGTAAGAAA 2977

QY      137 ----- 137
DB      2978 CAGCACTACATATCTAATAAAGTCTAATAATGATTAATAAATAAGTGTGAAGCAA 3037

QY      137 ----- 137
DB      3038 CAGTTATCCCTACAAAACTTTTGGGGTAGATCATCCAGAGGTTGTTCCAGACTCT 3097

QY      137 ----- 137
DB      3098 TGCTTAATAAATGTTTCTTAAATAAGTTTGAAGAAATGTTATATGATGAATAATAT 3157

QY      137 ----- 137
DB      3158 GAAAGAAAACATATCATATTTAATAAATAATAAGTATCAAGTAAGAAATTAACAATA 3217

QY      137 ----- 137
DB      3218 GGAATATATCTAATAAGAAATTTAGTGCTTTTCGTTAACAATATCTTATGTTATTC 3277

QY      138 -----1lePheGlyIle 141
DB      3278 ATTTGTTCTTTAATTTCCCTTCTTATTTTGTGAATTTACTAATGCAAGATTTTGGATC 3337

QY      141 IleuGlnIleSerSerIleSerAspIleValIleAspIleValIleValIleValIleVal 161
DB      3338 CTTAGAACATCTAATAAATTTTCACTACAAAGATGAAGGCTTTGGATTTGCTTTGA 3397

QY      161 pSerValGlnIleAsnLeuValIleGlyGlyGlyIleAsnGlyIleAsnGlyIleValIle 181
DB      3398 TAGTGTCAAAATTTATGTTGTGGAGGAGAGAACTATCATATGCAATGACCAAGATAG 3457

QY      181 pTrpProSerSerCysIleAsnIleSerIleu----- 192
```

Db 3458 GTGGCCAGGTTCTGGAAAATAAATACGTAATTTATTAACCTTGCTATTAAGT 3517  
QY 192 -----  
Db 3518 TTATACCTATGTTGTCGAAATCTTTAACTGTCTAAAGATATATATATTGAAGA 3577  
QY 192 -----  
Db 3578 GGTCACAAATGATCATATTTTAGAGATTCGACCAATATAGTTTATGTAATCTA 3637  
QY 193 -----Pr 193  
Db 3638 ATTTTCAGAGCATCTTGGCTTGTACTGATCATTTGTAACCTTTTCTTCATGCAGCC 3697  
QY 193 OCSaTPaSPAlaProthr-----  
Db 3698 ATCGAGGATGACCAACGGTACGTTAATGCAATTTGATTAATAAAAAAAGCCTA 3757  
QY 199 -----  
Db 3758 AAATATATTTGAATTTAATTGAAGTATATATATTTCTTAACCTTGGCAGAGCCTAT 3817  
QY 199 -----  
Db 3818 TACCCCTTGACCTAATTAATAGTATTTTAAGATATAAAAGTTTGTAGTGAACAA 3877  
QY 199 -----  
Db 3878 AATTAGATATTCAAAAAATTTGAAAATTTACTATATAATTCGAATTTTTCATATCA 3937  
QY 199 -----  
Db 3938 ATATGATTAAAAATATTAGTTAAAGTTCTTATGATTTGATCTGAAAAATAAATCAG 3997  
QY 199 -----  
Db 3998 ACAAAACAATAGTAGCGAGAAAGTATATACATACCTTCGAGTAGATCGATTGT 4057  
QY 199 -----  
Db 4058 ACAACACCTCAAAACCTACGTTTCTTGATTTATATTTCCATTTCTTTAATAGTAA 4117  
QY 199 -----  
Db 4118 TCAAGGCTATTGTTCTGCAAAATCTATACATTGGAACCTATCTTGACGCTGCT 4177  
QY 199 -----  
Db 4178 ACATTCGAGATCGTGAACAATGATGATATTTAACTTGATTTAAATTTAA 4237  
QY 200 -----AlaLeuThrPheTTPaenCylysaenLeuIysValAs 212  
Db 4238 ACTAATATTGTTAATTTTCAGGCTTAACTTCGGAATTCAGAAAATTTGAAGTGA 4297  
QY 212 nAsnLeuIysSerIysAsnAlaGlnGlnIleHisIleIysPheGlySerCythrAsnVa 232  
Db 4298 TAATCTAAAGAGTAAAAATGACACAAATTCATATCAATTTGATCATGCACCTAAGT 4357  
QY 232 IValIaSerAsnLeuMetIleAsnAlaSerAlaIysSerProAsnThrAspGlyValHi 252  
Db 4358 TGTAGCTTCAAAATTTGATGATCAATGCTTCAGCAAAAGAGCCAAATACTGATGAGTCA 4417  
QY 252 sValSerAsnThrGlnIyrIleGlnIleSerAspThrIleIleGlyThr----- 268  
Db 4418 AGTACAAATACTCAATATATTTCAAAATCTGATATCTATTAATGGAACAGGTTATTAAT 4477  
QY 268 -----  
Db 4478 TTAATTTTATTTATCAATTTAATTAGAAAAAAGAGATATTTTATTTGATTAACTA 4537  
QY 269 -----GlyAspAspCysIleSerIleValSerGlyse 279

Db 4538 AATTATTAATTTTAAATTTTTTTTATAGTGATGATGTATTTCAATTTCTTGATTC 4597  
QY 279 rGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHiIleGlyIleSer----- 296  
Db 4598 TCAAAATGTCAGGCGCACAAATATTTACTTGTGTCAGGTCATGATATAG-GTACTCTA 4656  
QY 296 -----  
Db 4657 TTTTACAAATATATCTGTTTCCATTTTCTATTTTCATTAAGAGTATGATATATATA 4716  
QY 296 -----  
Db 4717 TTACTTAAATCCTTAATTAATTAATTTATGCAAAATTTTCTCTTGCTTAATGTTAAT 4776  
QY 296 -----  
Db 4777 GACTTAGACAAATAATTAGGCGCGTTGATGGCGGAATAAAGCAGCTTTAAAAAGTA 4836  
QY 296 -----  
Db 4837 CTTTAAAAAGTGTGAACCTTATTTTAAATTAAGCATTAATCGGTTTGATTAAGTGC 4896  
QY 296 -----  
Db 4897 TGAAGTTTATGTCAAACGTGAAGGAAAAATGAGAGAAAGAAATGTAGGTTATA 4956  
QY 296 -----  
Db 4957 TGGGTTATTTGTATAAAAATATTAAAGCACAAAAAAGATAAAAATGTGTCAACTTAAAAA 5016  
QY 296 -----  
Db 5017 ACTATTAAGTACCTTACCTTACCCAGCTTTTAACTTTTGCTTAAATAAGTTTTT 5076  
QY 296 -----  
Db 5077 TTTTAAAACTTAAATAAATAGTTGTTTGTAGTATTCAGTAAAGCTAATAATGCAAAAACC 5136  
QY 296 -----  
Db 5137 AGCTTTAAGTCAGTTTGACACAGCTTTTAAAGTGCAGCAACAGAGCTTTAAATGTCG 5196  
QY 296 -----  
Db 5197 CTTAGATGTCATATATATTTGAGCTTTTGAAGTAGTATATATTCCTTAAGTCA 5256  
QY 296 -----  
Db 5257 CATTAATAACATGCTTTAACAATAGCACATATATGTTATCAAAAGCAAAATGATGATTA 5316  
QY 296 -----  
Db 5317 TTTTGGCAATTTGATTAATTCACAGAAAAAGGATAGTTCMAAGTATCAATTCATGAT 5376  
QY 296 -----  
Db 5377 TGAAGATATCATTAAGACTAAAAATTAGAAGATCAATAATTGAGGATCAAAAAATGTTAT 5436  
QY 296 -----  
Db 5437 TACCTTAATTAATAATCTATTCATTTCAATTAATTAATTAAGAGTGTATATA 5496  
QY 296 -----  
Db 5497 ATCTAATAAACAATGCAATTAATTAATGACGAATGTGTTTGGTACTTAATCTTCT 5556  
QY 297 -----IleGlySerLeuG1 301  
Db 5557 GAATATTTGCTATTTTCTCTTTTAAATTTTTCATGATTAATTTGAGAACCTTAGG 5616  
QY 301 IYserGlyAsnSerGlyAlaIyrIleValSerAsnValThrValAsnGlnAlaIysIleIleG1 321  
Db 5617 ATCTGAAATTCAGAGCTTAATGTCATTAATGTATTAAGTAAAGCAAAATTAATCGG 5676

```

QY 321 YAAAGLubenglyValArgIlelyThrTrpGln----- 332
Db 5677 TCCCGAAAATGAGTAGATCATAGACCTTGCA-GGTACCTCCCCCCCCCCCCCCCC 5735
QY 332 ----- 332
Db 5736 CAGAGCCCATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAGATTAA 5795
QY 332 ----- 332
Db 5796 ATTGATTTTATGTTGAATTTATATTTGATTAAGATGATTTTACTAGCTTCTAT 5855
QY 332 ----- 332
Db 5856 GTTATATGAAAAAATGTTGAGAACTTCAATTTATCTGCTGCTAAGTAAAT 5915
QY 332 ----- 332
Db 5916 GTGTTGCTTTGTTAGAAAGTTGGTTTATCCAGTTTGGGTGATGATTAACCAACTTA 5975
QY 332 ----- 332
Db 5976 TAAATGAAAAAGGGCTGCAACGGCCGCCACTAGTCTAGTATCATAGAAAGATCTAC 6035
QY 332 ----- 332
Db 6036 GTCTGTTTATTCAGATGAGCGTCTTGTTGATGTTAATTAATTAATTAATTAACA 6095
QY 333 ----- 338
Db 6096 TGTAAATTAAGCATTAATTAATTAATTAATGCTTAAATGAGGAGATCTGGAACAC 6155
QY 338 aSerAsnIlelyPheLeuAsnValGluMetGlnAspVallyrTyProIleIleIleAs 358
Db 6156 TAGCAACATCAAAATTTCTGAATGTGGAATGCAAGCGTTAAGTATCCCATATATTAA 6215
QY 358 pGlnAsnTyrcyAspAspValGluProcySileGln----- 370
Db 6216 CCAAAAATAATGTCGATCGATGGAACCATGATATACAAAGGTAAATTTTAAATTAACA 6275
QY 370 ----- 370
Db 6276 AATTTATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6335
QY 370 ----- 370
Db 6336 TTCACCTAGTAACTTTTGAATGATTTTGAATGATTTTGAATGATTTTGAATGATTT 6395
QY 370 ----- 370
Db 6396 TTTTCTTCTAATTTATGATCTTTTGAATGCTTAATAATCTGTTAAAGTAACT 6455
QY 370 ----- 370
Db 6456 GAATATCATAGAAAAATGTTAGATTAAATCTAAATTTTAAATTTTAAATTTTAAAGCTT 6515
QY 370 ----- 370
Db 6516 ATCTAATATTTTGTATGTAAGATGGAACATATATCAATTAACATTAGATTCAATTTA 6575
QY 370 ----- 370
Db 6576 ATAAATATCTAAATATTTGATGCAATCAATCATGACTACAAACGAATATACAGATT 6635
QY 370 ----- 370
Db 6636 CTCAAACATATATGATGAAGTCAATTTCAAAACGAATATATATGATATATCTTAA 6695
QY 370 ----- 370
Db 6696 AAGAGCATTTGGTAAATTAAGTAAATAATCATTAAGTTAATAAAATTTCTAACTGA 6755

```

```

QY 371 -----GlnPheSerAlaValGlnValyAsnVal 381
Db 6756 TCTTCACGATTATTAATCACTTTGTCACATTTTCAGACATTCAGGAAAAATGTG 6815
QY 381 aTyrcGluAsnIlelySerAlaThrLyValAlaIlelyPheAspCySer 401
Db 6816 TGTATGAGAAATTCAGAGCCACAGTGCACAAAGGTGCCATTAATTTGATTGACGA 6875
QY 401 hAsnProcySgllyIleIleMetGluAsnIleAsnLeuValGlyGluSerGly 421
Db 6876 CAACCTTTCATGATAGGAATTAATGAGAAATTAATTAATTTAGTAGGGAAGTGA 6935
QY 421 yProSerGluAlaThrCylAsnValHisPheAsnAsnIleGluHisValThrProH 441
Db 6936 AACCATCAGAGGCTACGTCGCAAAAATGTCATTTTAAATGATGTAACATGTTAC 6995
QY 441 lCySerThrSerLeuGluIleSerGluAspGluAlaLeuLeuTyrcyAsnTy 457
Db 6996 ACTGCATTCATAGAAATTTCAAGAGATGAAGTCTTTTGTATATTAAT 7045

RESULT 5
US-10-739-930-530
; Sequence 530, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; NUMBER OF SEQ ID NOS: 2003-12-18
; SEQ ID NO 530
; LENGTH: 1621
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER21779_1
US-10-739-930-530

Alignment Scores:
Pred. No.: 2,866-120 Length: 1621
Score: 1127.00 Matches: 223
Percent Similarity: 65.33% Conservative: 71
Best Local Similarity: 49.56% Mismatches: 116
Query Match: 47.15% Indels: 40
DB: 18 Gaps: 6

US-10-691-374-2 (1-457) x US-10-739-930-530 (1-1621)
QY 7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
Db 124 GCTGTTTCTTANGGCTCTTTTGAATGCTCGTGTGTGAAGCTTTGAGTAGCAAGCTT 183
QY 27 IleAsp-----AspAsnLeuPheLyGlnVal 35
Db 184 GATGATGATATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243
QY 36 TyrAspAsnIleLeuGluGlnGluPheAlaHisAspPheGlnAlaTyrcySerTyrcy 55
Db 244 AACGATGATGCTT-----AGCTTGAATAGCTGACG 276
QY 56 SerLyAsnIleGluSerAsnAsnAsnIleAspLyValAspLyAsnGlyIleLyVal 75
Db 277 GAACCACTTGGAGACATCA-----ACC 300
QY 76 IleAsnValLeuSerPheGlyAlaLyserGlyAspGlyLyserThrTyrcyAspAsnIleAlaPhe 95
Db 301 GTTAGGTTCACAACTTCGAGCCCAAGAGATGAGAAAAATGATGATGATGATGATGATGAT 360
QY 96 GluGlnAlaTyrcyAsnGluAlaCysSerSerAlaGlyThrProValGlnPheValValProLy 115
Db 361 AAGAAAGCATGGAAGAAAGCATGTTCAACAATGAGATTACTTCTTGGTCTCTAA 420

```

```

QY 116 AsnIysAsnTyrLeuLeuIysGlnIleThrPheSerGlyProCysArgSerIleSer 135
Db 421 GGAAGAGCTATCTCTTAAGTCTACTCGATTGAGGCCCAAGCAATCTTACGTAC 480
QY 136 ValIysIlePheGlySerLeuGlnIleAspSerIleSerAspTyrIysAspArg 155
Db 481 TTTCAGATCCTTACGATCTTATCAGATCTACGAAACGTTCAATTTACAAAGCAAAAC 540
QY 156 LeuTyrIleAlaPheAspSerValGlnAsnLeuValGlyGly-----GlyThr 173
Db 541 CATTGCTTATCTTAAAGAGCGTTAACTATCATCATGACGAGGCTTCACGCGGAAAT 600
QY 174 ILeAsnGlyAsnGlyGlnValTyrTrpProSerSerCysValIleAsnIysSerLeuPro 193
Db 601 ATTAAATGGCAACGAAACAACTGGTGGCAAACTCATGCAAAATGCAAAATCTAAAGCA 660
QY 194 CysArgAspAlaProThrAlaLeuThrPheTyrPasnCysIysAsnLeuValAsnAsn 213
Db 661 TGCACAAAGCTCCACCGGCTCTTACTTATCAATTTAAAGAAATTTGAATGTAGAAAT 720
QY 214 LeuIysSerIysAsnAlaGlnGlnIleHisIleIysPheGlySerCysThrAsnVal 233
Db 721 CTGAGGGGTGAAAATGCGCAGCATTCAGATTTCATTTGAGAAATGCAACAAAGTTGAA 780
QY 234 AlaSerAsnLeuMetIleAsnAlaSerAlaIysSerProAsnThrAspGlyValHisVal 253
Db 781 GTTATGTAATGTTGAGATCATCTGCGGCGATAGTCCACACAGATGATCATATC 840
QY 254 SerAsnThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIle 273
Db 841 ACTAATACTCAAAACATTCGAGTCTCCAACTCAATATCGGAACAGGTGATGATTGATA 900
QY 274 SerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHis 293
Db 901 TCATATGAGATGAGCAACGCAAAATCTTCAATCTTTGATTTAACTTGGCCCCGGTAC 960
QY 294 GlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsnValThr 313
Db 961 GGGATCGACATTTGGGAGCTTGGGGAGCAGACAAATTCGAAACCTTATGCTCGGAATTAAT 1020
QY 314 ValAsnGlnAlaIysIleIleGlyAlaGlnAsnGlyValArgIleIysThrTrpGlnGly 333
Db 1021 GTGATGTTGCTTAAGTTCTCTGAGAGTCAATGAGATTGAGATTAAAGCTTATACGGGA 1080
QY 334 GlySerGlyGlnIleAspSerAsnIleIysPheLeuAsnValGlnMetGlnAspValIysTyr 353
Db 1081 GGATCGAGAACTCCCAAGAACATTTAAATTTCAAAATATTCGATGGAATAACGTTCAAGAT 1140
QY 354 ProIleIleIleAspGlnAsnTyrCysAspArgValGlnProCysIleGlnGlnPheSer 373
Db 1141 CCGATCATATATCCAGCAGACGATCTGCGACAG-----GACAAATGCGAAGACCAAGATCG 1197
QY 374 AlaValGlnValIysAsnValValTyrGlnAsnIleIysGlyThrSerAlaThrIysVal 393
Db 1198 GCGATCGAAGTGAAGAAACGTTGTGTACAGAAACATATCTGGTACAGACGCTACCGAGTGT 1257
QY 394 AlaIleIysPheAspCysSerThrAsnPheProCysGlnGlyIleIleMetGlnAsnIle 413
Db 1258 GCGATTAACGTTGAATTCGAGCAGAAAGTATCATATCTCAAGGATTTGTCTTGAAACGCT 1317
QY 414 AsnLeuValGlyIysSerGlyIysProSerGlnAlaThrCysIysAsnValHisPheAsn 433
Db 1318 AAAATTAAGGA-----GAAACAGCTTCTGCAAAATGCGCAATGTTAAA 1362
QY 434 AsnAlaGlnHisValThrProHisCysThr 443
Db 1363 AATCAAGCAGCGTTTCTCTAAATGCTCT 1392

```

RESULT 6  
 US-10-151-668-1  
 : Sequence 1, Application US/10151668  
 : Publication No. US2002018460A1

```

: GENERAL INFORMATION:
: APPLICANT: UVSKOV, Peter
: APPLICANT: CHLD, Robin
: APPLICANT: VAN ONCKELIN, Henri
: APPLICANT: PRINSEN, Els
: APPLICANT: BORKHARDT, Bernard
: APPLICANT: SANDER, Lijl
: APPLICANT: PETERSEN, Morten
: APPLICANT: BUNDGARD, POUSEN, Gert
: APPLICANT: BOTTERMAN, Johan
: TITLE OR INVENTION: Seed Shattering
: FILE REFERENCE: 2121-0138P
: CURRENT APPLICATION NUMBER: US/10/151,668
: CURRENT FILING DATE: 2002-05-21
: PRIOR APPLICATION NUMBER: US/09/051,239
: PRIOR FILING DATE: 1998-09-28
: PRIOR APPLICATION NUMBER: PCT/EP96/04313
: PRIOR FILING DATE: 1996-10-04
: PRIOR APPLICATION NUMBER: EP 95 402241.4
: PRIOR FILING DATE: 1995-10-06
: PRIOR APPLICATION NUMBER: EP 95 203328.0
: PRIOR FILING DATE: 1995-12-08
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: Patencin Ver. 2.1
: SEQ ID NO 1
: LENGTH: 1631
: TYPE: DNA
: ORGANISM: Brassica napus
: FEATURE:
: OTHER INFORMATION: Location 95-163 = region encoding the presumed
: OTHER INFORMATION: endo-PG signal peptide.
: OTHER INFORMATION: Location 884-900 = region of the endo-PG cDNA
: OTHER INFORMATION: corresponding to oligonucleotide PG3
: OTHER INFORMATION: Location 1059-1073 = region of the endo-PG cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (95)-(1393)
: OTHER INFORMATION: Location 1229-1245 = region of the endo-PG cDNA
: OTHER INFORMATION: complementary to oligonucleotide PG5
: OTHER INFORMATION: Location 821-837 = region of endo-PG cDNA
: OTHER INFORMATION: corresponding to oligonucleotide PGL1.
: NAME/KEY: unsure
: LOCATION: (1439)
: OTHER INFORMATION:
: OTHER INFORMATION: Strain cv. Topaz.
: US-10-151-668-1
:
: Alignment Scores:
: Pred. No.: 1,23e-118
: Score: 1113.00
: Percent Similarity: 65.32%
: Best Local Similarity: 49.89%
: Query Match: 46.57%
: DB: 13
: Gaps: 8
:
: US-10-691-374-2 (1-457) x US-10-151-668-1 (1-1631)
:
: QY 7 SerIleLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
: Db 119 GCTGTTTCTTATGGCTCTTTTGTATGCTGCTGTCGCAAGCTTGAAGTACGAACTGA 178
:
: QY 27 ILeAspAsnLeuPheIysGlnValTyrAspAsnIleLeuGlnGlnIlePheAlaHis 46
: Db 179 ---GATGATGGA-----TATGTCAT 196
:
: QY 47 AspPheGlnAlaTyrLeuSerTyrLeuSerIysAsnIleGlySerAsnAsnAsnIleAsp 66
: Db 197 GAAAGATGGA-----AGCTTCGAATCCGATGTTATCAAGTCAACGACGACGAC 247
:
: QY 67 LysVal-----AspLysAsnGlyIleLysValIleAsnVal 78
:

```



```
Db 248 GACGTTCTTACCTTAAAAAGCTCTGATAGACCCACTACCGAATCATCTAGTGT 307
Qy 79 LeuSerPheGlyAlaIleGlyGlyAspGlyThrTrpAspAsnIleAlaPheGluGlnAla 98
Db 308 TCGAATCTCGAGCCAAAGAGATGAAAAACGATGATACCTCAGGCTTCAAGAAAGCA 367
Qy 99 TrpAsnGluAlaCysSerSerArgThrProValGlnPheValValProIleAsnIleAsn 118
Db 368 TGGAAAGAAAGCGATGTTCAACAAATGAGAGTACTACTACTTCTTAATCTTAAGAAAGACT 427
Qy 119 TyrIleuLeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIle 138
Db 428 TATCTCTTAATCTTAATGATTCAGAGCCCATGCAATCTTACGTACGTCCAGATC 487
Qy 139 PheGlySerLeuGluAlaSerSerLysLysIleSerAspTyr---LysAspArgLeuTrp 157
Db 488 CTAGGCACTTATTCAGCTTCTACAAACGATCGAATTACAGTAAGACAGAAACCACTGG 547
Qy 158 IleAlaPheAspSerValGlnAsnLeuValValGlyGlyGly-----GlyThrIleAsn 175
Db 548 CTATATTTGGAAAGCGCTTAATATATCTATCATCATGATGGCGGCTCGCGGAGATTGTGAT 607
Qy 176 GlyAsnGlyGlnValTrpTrpProSerSerCysLysIleAsnLysSerIleuProCysArg 195
Db 608 GCGAAGCGAAATATCTGTCGCAAAACTATCCAAATGACAAATCTTAAGCCATGACACA 667
Qy 196 AspAlaProThrAlaLeuThrPheTrpAsnCysLysAsnLeuLysValAsnAsnLeuLys 215
Db 668 AAGGCGCCAGCGCTTACTACTCTACCAACCTTAAGAAATTTGAATGTAAGATCTGAGA 727
Qy 216 SerLysAsnAlaGlnGlnIleIleIleLysPheGlySerCysThrAsnValValAlaSer 235
Db 728 GTGAAATAATGCACACAGATTCAGATTTCGATTGGAATAATGCACAACTTTCGCGCTTAA 787
Qy 236 AsnLeuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValIleValSerAsn 255
Db 788 AATGTTAAGATCACTGCTCTCGCGATGATCCCAACAGGATGGATTTATCTGCTGT 847
Qy 256 ThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIle 275
Db 848 ACTAAACCATTCGATCTCCAAATTCAGACATTTGGGACGATGATGATATTCATT 907
Qy 276 ValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyIleGlyIle 295
Db 908 GAGATGATCCCAAAATGTTCAATTCATGATTAACTTTCGCGCCCTCGATGCGAATC 967
Qy 296 SerIleGlySerLeuGlySerGlyAsnSerGluAlaTyrValSerAsnValThrValAsn 315
Db 968 AGCATTTGAGAGCTTGGGGATGACATTCCAAGCTTATGCGGAATGATGTGAT 1027
Qy 316 GluAlaLysIleIleGlyAlaGlnAsnGlyValArgIleLysThrTrpGlnGlyLysSer 335
Db 1028 GGTGCTAGCGCTCTCGAGCTGACATGACATGAGTAAAGATCAAGACTTACCAAGGAGGTCA 1087
Qy 336 GlyGlnAlaSerAsnIleLysPheLeuAsnValGlnMetGlnAspValLysTyrProIle 355
Db 1088 GGAATCTGCTTAAGAAATTTAAATTTCCAAACATTCGTATGATATGTCAAAGATCGAATC 1147
Qy 356 IleIleAspGlnAsnTyrCysAspArgValGlnProCysIleGlnGlnPheSerAlaVal 375
Db 1148 ATTAATCGACCAAACTACTGCAAAAG---GACAAATGCGAAACAGCAAAATCTGCGGTT 1204
Qy 376 GlnValLysAsnValValTyrGlnAsnIleLysGlyThrSerAlaThrLysValAlaIle 395
Db 1205 CAAGTGAACATGTCGTATATCAGAACATTAAGGTAACAGCGCAACAGATGCGGAGATA 1264
Qy 396 LysPheAspCysSerThrAsnPheProCysGlnGlyIleIleMetGlnAsnIleAsnLeu 415
Db 1265 ATGTTAATTTGAGTGAATATATCCATTCCAAGGATTTGCTGTTGATGATGAAACATC 1324
Qy 416 ValGlyGlnSerGlyLysProSerGlnAlaThrCysLysAsnValIleIleAsnAsnAla 435
Db 1325 AAGAGA-----GAAAAAGCTTCTGCGAAATATCTCAATGTTAAGATTA 1369
```

```
Qy 436 GluHisValThrProHisCys 442
Db 1370 GGCACCTGTTCTCTTAATATGC 1390

RESULT 7
US-10-787-958-31
/ Sequence 31, Application US/10787958
/ Publication No. US20040154053A1
/ GENERAL INFORMATION:
/ APPLICANT: Wyatt, Paul
/ APPLICANT: Roberts, Jeremy A.
/ APPLICANT: Whitelaw, Catherine
/ TITLE OF INVENTION: Signal Transduction Protein Involved in Plant Defence
/ FILE REFERENCE: 0623.0890000
/ CURRENT APPLICATION NUMBER: US/10/787,958
/ PRIOR FILING DATE: 2004-02-27
/ PRIOR APPLICATION NUMBER: GB9806113.8
/ NUMBER OF SEQ ID NOS: 38
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 31
/ LENGTH: 1657
/ TYPE: DNA
/ ORGANISM: Brassica napus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (145)..(1443)
US-10-787-958-31

Alignment Scores:
Pred. No.: 1,856-117 Length: 1657
Score: 1103.00 Matches: 221
Percent Similarity: 65.54% Conservative: 70
Best Local Similarity: 49.77% Mismatches: 125
Query Match: 46.15% Indels: 28
DB: 7 Gaps: 7

US-10-691-374-2 (1-457) x US-10-787-958-31 (1-1657)
Qy 7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
Db 169 GGTATTTTCTTATGCGTTCTTTTATGTCGCTGCTGCGCAAGCTTGAAGTGAACGTA 228
Qy 27 IleAsp-----AspAsnLeuPheLysGlnValTyrAspAsnIleLeuGlu 41
Db 229 GATGATGATATGTCATGAAAGATGAGAGCTTTCGAAC-----GATAGTTAATCAAG 282
Qy 42 GlnGlnPheAlaHisAspPheGlnAlaTyrIleuSerTyrLeuSerLysAsnIleGlySer 61
Db 283 CTCAAACGACGACGACGTTCTTACTTGAAGCTCCGATGACCCACTACCGAATCA 342
Qy 343 TCA-----ACGTGATGCTTTCGAATCTC 366
Db 366 TCA-----ACGTGATGCTTTCGAATCTC 366
Qy 366 TCA-----ACGTGATGCTTTCGAATCTC 366
Db 366 TCA-----ACGTGATGCTTTCGAATCTC 366
Qy 82 GluAlaLysGlyAspGlyLysThrTrpAspAsnIleAlaPheGlnGlnAlaTrpAsnGlu 101
Db 367 GGAGCAAAAGGTGATGAAAAACCGATGATATCTCAGGCTTTCAGAAACATGCAAGAA 426
Qy 102 AlaCysSerSerArgThrProValGlnPheValValProIleAsnIleAsnIleuLeu 121
Db 427 GCATGTTCAACAAATGAGTGACTACTTCTTGAATCTTAAAGGGAAGACTTATCTCTT 486
Qy 122 LysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGlySer 141
Db 487 AAGCTATTAATGATTCAGAGGCCCATGCAAAATCATTAACGATTCAGATCCAGGCACT 546
Qy 142 LeuGluAlaSerSerLysIleSerAspTyr---LysAspArgLeuTrpIleAlaPhe 160
Db 547 TTATAGCTTTCACAAACGATCGGATTAACGTAATGACAAAGAACCACTGCTATTTTG 606
Qy 161 AspSerValGlnAsnLeuValValGlyGlyGly-----GlyThrIleAsnGlyAsnGly 178
```

```

Db      607 GAGACCTTAATTAATCTATCAATCGATGCGCGCTCGCGCGCTGATGTTATGAGCAACGGA
Qy      179 GlnValTrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspIlePro
Db      667 AAATATGGTGGCAAAATCATCATCAAAATGCAAAATCTAAGCAGTCACAAAGCGCCA
Qy      199 ThrAlaLeuThrPheTrpAsnLysValAsnLeuLysValAsnLeuLysSerLysAsn
Db      727 ACGGCTTCTTCTCTTACCACTTAAACAAATGTAATGTAAGAACTGAGAGTGAAGAAAT
Qy      219 AlGlnGlnIleHisIleLysPheGluSerCysThrAsnValValAlaSerAsnLeuMet
Db      787 GCACACGAGATTCAGATTGCAATGGAATGCAACAGCTTGATGATTAAAGATTAG
Qy      239 IleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTrp
Db      847 ATCACTGCTCTCGGCATATGTCACACAGATGGATTCATATGCTGCTACTATAAAG
Qy      259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly
Db      907 ATTCGAATCTCAATTCAGACATTCGACAGCTGATGATTCATATTCATTCATGAGGATGA
Qy      279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGly
Db      967 TCGCAAAATGTTCAATCAATGATTTAACTTGGCGGCCCGCTCATGGCATCAGATTGGA
Qy      299 SerLeuGlySerGlyAsnSerGlnAlaThrValSerAsnValThrValAsnGlnAlaLys
Db      1027 AGCTTGGGGATGACAAATTCGAAAGCTTATGTAATGGAATTAATGGAATGGTGTACG
Qy      319 IleIleGlyAlaGlnAsnGlyValArgIleLysThrTrpGlnGlySerGlyGlnAla
Db      1087 CTCTCTGAGCTCAATGAGTAAGATCAAGACTTACACAGGAGGCTCAGAACTGCT
Qy      339 SerAsnIleLysPheLeuAsnValGluMetGlnAspValLysTrpIleIleLeuAsp
Db      1147 AAGAACATTAATTCGAAACATTCCTGATGATGATTAATCAAGATCCGATCAATATCGAC
Qy      359 GlnAsnTrpCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValLys
Db      1207 CAGAACCTACGCGCAAG--GACCAATGCGAACACAGAAATCTCGGCTCAAGTGAAC
Qy      379 AsnValValTrpGlnAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAsp
Db      1264 AAGTGTGTATCGGAACATCAAGGTAACGCGCAACGATGTGCGTAAATGTTTAAAT
Qy      399 CysSerThrAsnPheProCysGlnGlyIleIleMetGlnAsnIleAsnLeuValGlnLys
Db      1324 TGCAGGTGTAATTCATGCGCAAGGTAATGCTTGAGATGTAACATCAAGGA--
Qy      419 SerGlyLysProSerGlnAlaThrCysLysAsnValHisPheAsnAsnAlaGlnHisVal
Db      1381 -----GGAAGAGCTTCTTGCAGAAATGTCAGATTAAGATTAAGACGCGTT
Qy      439 ThrProHisCys 442
Db      1429 TCTCTTAATGC 1440

```

```

RESULT 8
US-10-739-930-1548
; Sequence 1548, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739, 930
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 1548
; LENGTH: 1707

```

```

; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: Clone ID: BRANA-23APR03-CLUSTER5_1
US-10-739-930-1548

Alignment Scores:
Pred. No.: 1,95e-117 Length: 1707
Score: 1103.00 Matches: 221
Percent Similarity: 65.54% Conservative: 70
Best Local Similarity: 49.77% Mismatches: 125
Query Match: 46.15% Indels: 28
DB: 18 Gaps: 7

```

```

US-10-691-374-2 (1-457) x US-10-739-930-1548 (1-1707)
Qy      7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal
Db      169 GCTATTTCTTATGCGTTCTTTTATGCTCGCTGCGCAAGCTTGAGTGAAGCAAGTA
Qy      27 IleAsp-----AspAsnLeuPheLysGlnValTrpAspAsnIleLeuGlu
Db      229 GATGATGATATGTCATGAAGATGAAGCTTCGAAACC-----GATAGTTATATCAG
Qy      42 GlnGluPheAlaHisAspPheGlnAlaTrpLeuSerTrpLeuSerLysAsnIleGluSer
Db      283 CTCACACACGACGACGACGCTTACCTTGAAAAGCTCCGATAGACCCACTACCGAATCA
Qy      62 AsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSerPhe
Db      343 TCA-----ACTGTTAGTGTTCGAACTTC
Qy      82 GlnAlaLysGlyAspGlyLysThrTrpAspAsnIleAlaPheGlnGlnAlaTrpAsnGlu
Db      367 GAGACCAAAAGGTGATGAAAAACCGATGATACACAGCTTCAAGAAAGCATGAAAGAG
Qy      102 AlaCysSerSerArgThrProValGlnPheValValProLysAsnLysAsnTrpLeuLeu
Db      427 GCATGTTCAACAAATGAGAGTACTTCTTCTTATCTTAAAGGAAAGCACTTATCTCCTT
Qy      122 LysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGlySer
Db      487 AAGTCATTAATGATTCAGAGGCCATCAATCAATTCATGAGCTTCCAGATCCTTACGCACT
Qy      142 LeuGlnAlaSerSerLysIleSerAspTrp---LysAspArgArgLeuTrpIleAlaPhe
Db      547 TTATCAGCTTCTTCAAAAGATCGGATTCAGTAAATGACAAAGACCACTGCTTATTTTG
Qy      161 AspSerValGlnAsnLeuValValGlyGlyGly-----GlyThrIleAsnGlyAsnGly
Db      607 GAGGAGCTTAATTAATTAATCAATCATGCGCGCTCGCGGGGATTTGATGAGCAACGGA
Qy      179 GlnValTrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspIlePro
Db      667 AAATATGTTGCAAAATCATCATCAAAATGCAAAATCTAAGCAGTCACAAAGCGCCA
Qy      199 ThrAlaLeuThrPheTrpAsnLysValAsnLeuLysValAsnLeuLysSerLysAsn
Db      727 ACGGCTTCTTCTCTTACCACTTAAACAAATGTAATGTAAGAACTGAGAGTGAAGAAAT
Qy      219 AlGlnGlnIleHisIleLysPheGluSerCysThrAsnValValAlaSerAsnLeuMet
Db      787 GCACACGAGATTCAGATTGCAATGGAATGCAACAGCTTGATGATTAAAGATTAG
Qy      239 IleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTrp
Db      847 ATCACTGCTCTCGGCATATGTCACACAGATGGATTCATATGCTGCTACTATAAAG
Qy      259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly
Db      907 ATTCGAATCTCAATTCAGACATTCGACAGCTGATGATTCATATTCATTCATGAGGATGA
Qy      279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGly

```

```
Db 967 TCGCAAAATGTTCAATCAATGATTAACTTGGGCCCCGTCATCGCATCGCATCGATGGA 1026
Qy 299 SerLeuGlySerGlyAenSerGluAlaTyValSerAenValThrValangluAlaTyS 318
Db 1027 ACCTGGGGGATGACAAATTCCAAAGCTTATGTATCGGAATTAAATGTGATGGTGCTAG 1086
Qy 319 IleIleGlyAlaGluAenGlyValArgIleLeuThrTrpGlnGlyGlySerGlyGlnAla 338
Db 1087 CTCTGTGAGACTGACAAATGAGTAAGAAATCAAGACTTACAGGGAGAGGTCAAGAACTGCT 1146
Qy 339 SerAenIleLeuPheLeuAenValGluMetGlnAenValTySerProIleIleLeuAer 358
Db 1147 AAGAACATTAAATTCAAAACATCTCGTATGATTAATGTAAGATCCGATCTAATATCGAC 1206
Qy 359 GlnAenTyTyCyAenArgTyValGluProCyValIleGlnPheSerAlaValGlnValTyS 378
Db 1207 CAGAACTACTGGCAGCAAG---GACAAATCGAAACAAGAAATCTCGGTTCAAGTGAAC 1263
Qy 379 AenValValTyThrGluAenIleLeuSerAlaThrLeuValAlaIleLeuPheAer 398
Db 1264 AATGTGTGTATCGAAACATACAGGTGAGCGCAACGATGTGGCATTAATGTTTAAT 1323
Qy 399 CysSerThrAenPheProCysGluGlyIleIleMetGluAenIleAenLeuValGlyGlu 418
Db 1324 TGCAGTGTAAATATCATCGCAAGGTATGTGCTTGAGAAATGTGAACATCAAGGA--- 1380
Qy 419 SerGlyLeuProSerGluAlaThrCyValAenValIlePheAenAenIleGluHisVal 438
Db 1381 -----GAAAAGCTTCTTCAAAAATGTCAATGTTAAAGATTAAGGACCGTT 1428
Qy 439 ThrProHisCys 442
Db 1429 TCTCTTAATGC 1440

RESULT 9
US-10-437-963-24607
; Sequence 24607, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 24607
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29575C.1
US-10-437-963-24607

Alignment Scores:
Pred. No.: 5,55e-107 Length: 1182
Score: 1011.00 Matches: 187
Percent Similarity: 69.52% Conservative: 57
Best Local Similarity: 53.28% Mismatches: 105
Query Match: 42.30% Indels: 2
DB: 18 Gaps: 1

US-10-691-374-2 (1-457) x US-10-437-963-24607 (1-1182)
Qy 72 GlyIleLeuValIleAenValLeuSerPheGlyAlaLeuGlyAenGlyLeuThrTyThrAer 91
```

```
Db 79 GGCAGTAAATGTCTTACGACATACAGAGCTACGGGGCTCATGAGAGAGGACATGATGAC 138
Qy 92 AenIleAlaPheGluGlnAlaTrpAenGluAlaCySerSerAerTyThrProValGlnPhe 111
Db 139 ACCAAAGCATTTGGAGATPACATGGCTCGACCTTCTCTCTGCAAAACCTTCAGATTTCG 198
Qy 112 ValValProLeuAenValAenTyLeuLeuTyGlnIleThrPheSerGlyProCysArg 131
Db 199 CTCATCTCCCAAGGAGCAAGAAATACCTGATCAAGACACAAACACTGCTGTGTCACAGANA 258
Qy 132 SerSerIleSerAenValLeuPheGlySerLeuGluAlaSerSerTyIleSerAerTyTy 151
Db 259 TCAACATCTCTCATATGATGTGAAGGTATGTTGGTGGCTTCACAGAGAGTCAAGCTGG 318
Qy 152 TyAenArgTy-----ArgLeuTrpIleAlaPheAenSerValGlnAenLeuValGly 169
Db 319 ACAAAGAGACCATTAAGGACATGATTCGATCAGTGTGTCACTGTCTTACTGTCACT 378
Qy 170 GlyIleGlyTyThrIleAenGlyAenGlyGlnValTrpTrpProSerSerCysAenIleAen 189
Db 379 GGTGTGGGACCATAGATGAAATGCAAGATTGTGTGCAAAATTCATGCAAAACCAAC 438
Qy 190 TySerLeuProCysArgAenAlaProThrAlaLeuThrPheTrpAenCysAenLeu 209
Db 439 TCAAGCTTCCTCATGACACAGAGCTCCAAAGCGACATGACTTACTCTCGCAAGATCTG 498
Qy 210 TyValAenAenLeuLeuSerTyAenAenAenGlnIleIleIleTyPheGluSerCys 229
Db 499 AAGGTAGATATCTGAAGGTGTGAAACAGCCGAAATTCAGATTTCAAGTGAAGATTCG 558
Qy 230 ThrAenValValAlaSerAenLeuMetIleAenAlaSerAlaLeuSerProAenThrAer 249
Db 559 ACCGATGTATGTGTCTGCTGTCGATCACAGACCAAGAAACAGCCCAACACTGAT 618
Qy 250 GlyValIleValSerAenThrGlnTyIleGlnIleSerAerThrIleGlyTyArgly 269
Db 619 GGAATCTCATATCACACGACAGAGATGTGAGGTGACAGATGATGATCAAGACCGGG 678
Qy 270 AenAerCysIleSerIleValSerGlySerGlnAenValGlnAlaThrAenIleThrCys 289
Db 679 GATGACTGCATGTCAATCGAGGACGAAACCGAAGACTGCATGTCAAGAACTGTGTGT 738
Qy 290 GlyProGlyHisGlyIleSerIleGlySerLeuGlySerGlyAenSerGluAlaTyVal 309
Db 739 GAGCCGGGACAGGATCAGCATCGGTGCTGGGTGATCATTAATTCGAACTCATGTC 798
Qy 310 SerAenValThrValAenGluAlaTyIleIleGlyAlaGluAenGlyValArgIleTyS 329
Db 799 AACATATGTCACCTGTGACAAATGTCAAGTTTATGACACAGCCACGAGACTCGCATCAAG 858
Qy 330 ThrTrpGlnGlyGlySerGlyGlnAlaSerAenIleLeuPheLeuAenValGluMetGln 349
Db 859 ACATGCGAGGAGGAAAGGTTCCAGCAAGAACATCTGTTCCAGAACATGTGATCGATGAC 918
Qy 350 AenValTyTyProIleIleIlePheGlnAenTyTyAenArgTyValGluProCysIle 369
Db 919 AATGTCTGAAACCCATCATCATCAACAAACTACTGTGACTCTTCAACCTCGCAAG 978
Qy 370 GlnGlnPheSerAlaValGlnValTyAenAenValTyGluAenIleLeuGlyTyThrSer 389
Db 979 CAACAGAAATCTGCAAGTGAAGGTGAGCAATGTGCTTCAABAACATCAGGGGCAACAT 1038
Qy 390 AlaThrTyValAlaIleLeuPheAerCysSerThrAenPheProCysGluGlyIleIle 409
Db 1039 GCATCAGAGAGGACCATCATGCTGCATTCACACAGCATGTACTTGCCATGGCATTAACC 1098
Qy 410 MetGluAenIleAenLeuValGlyGluSerGly 420
Db 1099 TTGAGAAATGTCAATCTCACTGTCAAGGAGGA 1131

RESULT 10
US-10-437-963-24608
```



Db 202 GTTTCAGCCTGCAGACTACGGCGGCGACGGCGGAGCGGACGACGCGGCGG 261  
Qy 95 PhegluGlAlaTrpAenGluaIaCyseSerArThProValGlnPheValPro 114  
Db 262 CTGGCAGAGCGCGGAGCGCGCTGCTCGCGCGCGCGCGCTGCTGCTGCTCGCG 321  
Qy 115 LysAenLysAenTyrlEuLeuLysGlnIleThPheSerGlyProCyAaRgsSerIle 134  
Db 322 GAGAGCAGAGACTCTGCTGCTCGCGCAATCACCCTCTCGCGCGCTGCAATCCACATC 381  
Qy 135 SerValLysIlePheGlySerLeuGluaIaSerSerLysIleSerApyrLysAap--- 153  
Db 382 AGCTCATGTGAAGGAGGAGCGCTGCGGCTGCGCGGACATGTCGAATCGAAGCGAGAC 441  
Qy 154 ---AGATGLeuTrpIleAlaPheAapSerValGlnAenLysValGlyGlyGly 172  
Db 442 AACAGAGGTAAGTGCATGCTGCTCGCGCGGCTGCGCGGCTGCTGCGCGCGCGG 501  
Qy 173 ThrIleAenGlyAenGlyGlnValTrpTrpProSerSerCyAaLysIleAenLysSerLeu 192  
Db 502 ACCATGACGCGCAACGCGAGCGGCTGCGGAGACTCTGCAAGATCAACAGGCGCATC 561  
Qy 193 ProCyAaRgsApyrLysAaProThraIaLeuThrPheTrpAenCyAaLysAenLysValAa 212  
Db 562 CCATGCAAGGAGCTCCAAAGCGCGCTGAGTTTCACACGTCGACAACTGAGCGCGAAT 621  
Qy 213 AenLysSerLysAenAlaGlnGlnIleHisIleLysPheGlySerCyThraAenVal 232  
Db 622 GGTCTGAAGTGTGTAACAGCCGACGATCCATGTCAGTGAAGATTGCAACCGGCGG 681  
Qy 233 ValAlaSerAenLysIleAenAlaSerAlaLysSerProAenThraPheGlyValHis 252  
Db 682 GAGCTGCGCCATCTGCAATCTGCGCGCGCGGACGAGCCCAACACAGATGGCATCAC 741  
Qy 253 ValSerAenThrxGlnTrpIleGlnIleSerApyrThrIleGlyThrxGlyAaPaeCyA 272  
Db 742 ATCACCACGACGAAAGCGTCAAGTCAAGCATGCACTTAAAGACAGGAGTGACTGT 801  
Qy 273 IleSerIleValSerGlySerGlnAenValGlnAlaThraAenIleThrxGlyProGly 292  
Db 802 GTGTCATGAGAGATGGAGCCACGCGTTTACGTCACGAGATGCTGTGCGCGCGG 861  
Qy 293 HisGlyIleSerIleGlySerLeuGlySerGlyAaSerGlyLysValSerAenVal 312  
Db 862 CATGGATTTAGCATTTGAGCTTAGAGATGACAACTCCAGACTGAGCTGATGCATC 921  
Qy 313 ThrValAenGlyLysIleGlyAlaGlnAenGlyValAaGlyLeuThrxTrpGln 332  
Db 922 TTCATGACACCGCTGACCTCTATGGCACACCAATGGAGCTGGATCAAGACATGGCAG 981  
Qy 333 GlyGlySerGlyGlnAlaSerAenIleLysPheLeuAenValGlnAenValLys 352  
Db 982 GAGGAGAGTGAATACCCCAAGATATCGATTCCAAACATGCTCATGAACATGTCAG 1041  
Qy 353 TyrProIleIleIlePheGlnAenTyrcyAaApyrValGlnProcyA---IleGlnGln 371  
Db 1042 AACCCATATATCTTAACCAAACTACTGCGACTGACTGAAGAAAGCGAGACAGAGAG 1101  
Qy 372 PheSerAlaValGlnValLysAenValValTrpGlnAenIleLysGlyThrxSerAlaTr 391  
Db 1102 GATTCAGCAGTGAATGATGATGCTTCAAGAACATCCGACGAGACAAATTTGC 1161  
Qy 392 LysValAlaIleLysPheAapCySerThraPheProCyGlnGlyIleIleMetGln 411  
Db 1162 AAGAGTGCCTACTCTGAATGACAGAGACTACCAATGCTATGATATTTCTTACAG 1221  
Qy 412 AenIleAenLys-----ValGlyGlySerGlyLysPheSerGlyLysValAa 429  
Db 1222 GACATCACTCGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1278  
Qy 430 ValHisPheAenAenAlaGlnHisValThrxProHis---CyAaThrxSer 444  
Db 1279 GCMAAATGAGGAAATCTGGAAAGTTGTTCCAAACCAATGACACAGC 1326

RESULT 12  
US-10-260-238-618  
Sequence 618, Application US/10260238  
Publication No. US20040016025A1  
GENERAL INFORMATION:  
APPLICANT: Budworth, Paul R.  
APPLICANT: Moughamer, Todd G.  
APPLICANT: Briggs, Steven P.  
APPLICANT: Cooper, Bret  
APPLICANT: Glazebrook, Jane  
APPLICANT: Goff, Stephen A.  
APPLICANT: Katagiri, Fumiyaki  
APPLICANT: Kreps, Joel  
APPLICANT: Provart, Nicholas  
APPLICANT: Rieke, Darrell  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
FILE REFERENCE: 60111-NP  
CURRENT APPLICATION NUMBER: US/10/260,238  
CURRENT FILING DATE: 2002-09-26  
PRIOR APPLICATION NUMBER: US 60/325,448  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/325,277  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/370,620  
PRIOR FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 6077  
SEQ ID NO 618  
LENGTH: 1168  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: N\_region  
LOCATION: (6)..(8)  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N\_region  
LOCATION: (46)..(46)  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N\_region  
LOCATION: (49)..(49)  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N\_region  
LOCATION: (102)..(102)  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N\_region  
LOCATION: (190)..(190)  
OTHER INFORMATION: n = any nucleotide  
US-10-260-238-618  
Alignment Scores:  
Pred. No.: 2,18e-102 Length: 1168  
Score: 971.50 Matches: 185  
Percent Similarity: 67.82% Conservative: 70  
Best Local Similarity: 49.20% Mismatches: 114  
Query Match: 40.65% Indels: 7  
DB: 17 Gaps: 5  
US-10-691-374-2 (1-457) x US-10-260-238-618 (1-1168)  
Qy 75 ValIleAenValLysSerPheGlyAlaLysGlyAaPylLysThrxAenIleAla 94  
Db 35 GTTTTACGCTGAGAAAGTTACGGGGGAGACGGGACGGGACGACGACGCGCGG 94  
Qy 95 PhegluGlAlaTrpAenGluaIaCyseSerArThProValGlnPheValPro 114  
Db 95 CTGGCAGAGCGCGGAGCGCGCTGCTCGCGCGCGCGCTGCTGCTGCTCGCG 154  
Qy 115 LysAenLysAenTyrlEuLeuLysGlnIleThPheSerGlyProCyAaRgsSerIle 134

```
Db      155 GAGAGCAGGAGCTACCTGCTCCGCAAGTACCTTCCGCGCCCTGCAATCCACATC 214
Qy      135 SerVallysllePheglYserleuGlnAlaSerSerlylelSerAspTyrlyAsp--- 153
Db      215 AACCTAGGTGTGAAGGGAGCGCTGTGGCTCGCCGCAATGTGCAACTGCAACAGAGAC 274
Qy      154 ---ArgArgleuTrrlleAlaPheAspSerValGlnAsnleuValAlGlyGly 172
Db      275 AACAGAGGTACTGATCTGCTCCGCGGCTGAGAGGAGTCTGCGCTGCGCGCGCGC 334
Qy      173 ThrllAsnGlyAsnGlyGlnValTrrTPProSerSerCyAlGlylAsnlySerleu 192
Db      335 ACCATGACCGGCAACGGCGGTGTGGAGAACTCTGCAATCAACAGGCGCATCTC 394
Qy      193 ProCyArgAspAlaProThrAlaLeuThrPheTrrPAsnCylAsnleuLyValAsn 212
Db      395 CCATGCAAGGAGCTCCAAAGGCGCTGAGTTTCCACAGTGGACAACTGAGCGCTGAT 454
Qy      213 AsnleuLySerlyAsnAlaGlnGlnlleHlsllElysPheGluSerCyAsnVal 232
Db      455 GGTCTGAGATGTGTGAAGCCAGAGATCCACATGTACGTGAGAGATTGCAACGGGCTG 514
Qy      223 ValAlaSerAsnleuMetlleAsnAlaSerAlaLySerProAsnThrAspGlyValHls 252
Db      515 GAGCTGGCCCACTGTCAATCTGTGCGCCGCAAGAGCCCAACAGATGGCATCAC 574
Qy      253 ValSerAsnThrGlnTrrlleGlnlleSerAspTrrlleGlyThrGlyAspAspCy 272
Db      575 ATCAACCAACAGAAAGATCTCAAGTACAGCACTGCACATTAAAGACGGGAGATCTGT 634
Qy      273 lleBerlleValSerGlySerGlnAsnValGlnAlaThrAsnleuThrCySglProGly 292
Db      635 GGTGTGATGTAGAGATGGAGCCAGCTTTCACCTGCAAGACTGAGTGTGGGCGCGG 694
Qy      293 HlsGlylleSerlleGlySerleuGlySerGlyAsnSerGlnAlaTrrValSerAsnVal 312
Db      695 CATGGATTTACATTTGGAGCTTGTAGAGATGACAACTCCAGAGCTGAGGTGTGACATC 754
Qy      313 ThrValAsnGlnAlaLysllelleGlyAlaGlnAsnGlyValArglleYsThrTPGln 332
Db      755 TTCATTCACACCCGTGACCTCTATGAGCACCAACATGAGCTGGATCAAGATGCGAG 814
Qy      333 GlylySerGlyGlnAlaSerAsnlelleYsPheleuAsnValGlnMetGlnAspValys 352
Db      815 GAGAGGAGTGTATGCGCAAGATATGTATTCAGAAATGGTCAATGAAACAGTGTCAAG 874
Qy      353 TyrProillelleAspGlnAsnTrrCyAspAspArgValGlnProCys---lleGlnGln 371
Db      875 AACCAATTAATCATTTGACCAAACTACTGCGACTCAAGCTAAGAAATGCGAGACAGGAG 934
Qy      372 PheSerAlaValGlnVallyAsnValAlTrrGlnAsnlelleYsGlyThrSerAlaThr 391
Db      935 GGTACACAGATGTAGATTCAGAAATGTGCTTTCAAACATCGCAGGAGCAACAATTTCC 994
Qy      392 LyseValAlaIlelySerpheAspCySerTrrAsnPhProCySglGlyllelleMetGln 411
Db      995 AAGAGTGCCTACTCTGAACTGCACAGAACTACCATCTGTATGACATTTCTTACAG 1054
Qy      412 AsnleuAsnleu-----ValGlylySerGlyLySerpSerGlnAlaThrCySglAsn 429
Db      1055 GATATCAACCTGAAATGTGTGATGCAATGTGT---GCCACAGAAAGTACTTGCCAGAT 1111
Qy      430 ValHlsPheAsnAsnAlaGlnHlsValTrrProHls---CyeThrSer 444
Db      1112 GCMAAATGAGAAATCTGMAACAGTGTTCACAAACCATGACACAGC 1159
```

## RESULT 13

```
US-10-425-115-33752
; Sequence 33752, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21 (53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 33752
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_130788C.1
US-10-425-115-33752

Alignment Scores:
Pred. No.: 1,33e-98 Length: 1372
Score: 940.00 Matches: 174
Percent Similarity: 71.22% Conservative: 66
Best Local Similarity: 51.63% Mismatches: 93
Query Match: 39.33% Indels: 4
DB: 18 Gaps: 3

US-10-691-374-2 (1-457) x US-10-425-115-33752 (1-1372)

Qy      114 ProLyAsnLyAsnTrrlleuAlaPheAspSerValGlnleuThrPheSerGlyProCyArgSerSer 133
Db      12 CCGCGCGAGCGGTACTCTGCTGAGCTCTCAACCTCCGTGCGCGGTGCAAGTCCAGC 71
Qy      134 lleSerVallysllePheglYserleuGlnAlaSerSerlyslleSerAspTyr----- 151
Db      72 GTCAAGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 131
Qy      152 LyAspAspArgleuTrrlleAlaPheAspSerValGlnleuThrPheSerGlyProCyArgSerSer 171
Db      132 AATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 191
Qy      172 GlyThrllEAsnGlyAsnGlyGlnValTrrTPProSerSerCySgllylleAsnlySer 191
Db      192 GCGCCATTCATGTGCAACGCGCGAGAAATGTGTGCTTCATGCTGCAAGTCAACAAAGCT 251
Qy      192 LeuProCyArgAspAlaProThrAlaLeuThrPheTrrPAsnCylAsnleuLyVal 211
Db      252 CTCCTTGCAGAGAGGCTCCAGCGCTGTCAATTCACATGCTGCGCTGCAACCTGAAAGTG 311
Qy      212 AsnAsnleuLySerlyAsnAlaGlnGlnlleHlsllElysPheGluSerCySglThrAsn 231
Db      312 GAGAAATCTGAAATTTATGACAGCAACAGATCCACATGTCACTGAGAGACTTCGAAAT 371
Qy      232 ValAlaSerAsnleuMetlleAsnAlaSerAlaLySerProAsnThrAspGlyVal 251
Db      372 GTGCTGTGCGCGGTGTTCATCACAGCGCTGCAACAGCTTAACTGACCGCATC 431
Qy      252 HlsValSerAsnThrGlnTrrlleGlnlleSerAspTrrlleGlyThrGlyAspAsp 271
Db      432 CACATCACTGCGCAAGATGTAGCGCTCACAGCTGCAAGATCAAGACAGGAGCGAGC 491
Qy      272 CyeIleSerlleValSerGlySerGlnAsnValGlnAlaThrAsnlelleThrCySglPro 291
Db      492 TGCATGTCAATCGAAGCGGAGCTCAACCTTCATGTCTCAAGAGTACTTGCTGATCA 551
Qy      292 GlyHlsGlylleSerlleYserleuGlySerGlyAsnSerGlnAlaTrrValSerAsn 311
Db      552 GGGCATGGATATGATCGAGTCCAGGAGCTTGTAGAGCGACAACTCAAGAGCAAGAGTCTCGCGC 611
Qy      312 ValThrValAsnGlnAlaLysllelleGlyAlaGlnAsnGlyValArglleYsThrTrrP 331
Db      612 ATCAACCATGATTCAGTGCATCAACGCAACCAACGAGGAGCGCATCAAGAGCTATC 671
Qy      332 GlnGlylySerGlyGlnAlaSerAsnlelleYsPheleuAsnValGlnMetGlnAspVal 351
```

```
Db      672  CAGGAGGAGGAGGATACCGCAAGACATCAGCTTCCAAAATATGTTATGTACGACGTC 731
Qy      352  LVSTYRPIIlelleleAaRglnAaNTYTCyAaRPaRgValGluProCyv1leGlnGln 371
Db      732  AAGAACCCCATATCATCTCCGACCGAAGCTACTGCGACAGAGGCTAAGCCATTCGCGAAGAA 791
Qy      372  PhSeRAlValGlnValLyAaenValValTyrGlnAaenllelyseRlyThrsRAlaThr 391
Db      792  GAGTCGAGCATGAGGTTCAGCGGTGTCTTAAAGACATTAGAGGAGGACGACCACTACC 851
Qy      392  LVValAlAllelyeRheaRPySeSerThraenPhRProCyseGlnGlyllelleMetGln 411
Db      852  AAGAGCCCATCAAGATGAACTGACATGAGAGAGTCCCATGCAAGGCACTTACCTTCAG 911
Qy      412  AaenlleAaenlleuValGlyuseRgLyLyV--ProSeRglnAlaThRcyvlyAaenVal 430
Db      912  AAGATGACACTGAAATTCAGAGGACGCGCAAGGCGCAACAAGAGCACTTGTCAAGATGCA 971
Qy      431  HleRheaAaenlleAgluHleValThRProHle--CyethRseleuglu 446
Db      972  AAATGACAGAAATTTGGAGACATTCGCCCGGACGCTTGACATGCAATAAG 1022

RESULT 14
US-10-425-114-30261
/ Sequence 30261, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Sreen, Steven E
/ APPLICANT: Tabaka, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 30261
/ LENGTH: 1308
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZMFLB73002C03_FLI
US-10-425-114-30261

Alignment Scores:
Pred. No.: 1,766-90 Length: 1308
Score: 870.00 Matches: 162
Percent Similarity: 70.85% Conservative: 64
Best Local Similarity: 50.78% Mismatches: 89
Query Match: 36.40% Indels: 4
DB: 17 Gaps: 3

US-10-691-374-2 (1-457) x US-10-425-114-30261 (1-1308)
Qy      132  SeRseRlSeRValLyAllePhRgLySeRleugluAlaSeRseRlySleSeRAlaPyr 151
Db      2  TTCAGGTCACGCTCAGCGTGAAGGCGACCTGTTGCATCGCCAAACCGGCGGATGG 61
Qy      152  -----LyAaRPaRgRgLeuTPrilleAaPhAaRSeRValGlnAaenValValGly 169
Db      62  ACGGCAATGACAGGAGGACCTGGATCGTGTCCGAGCATTCACAACTCACCGCTCAAC 121
Qy      170  GLyGlyGlyThRlleaenGlyAaenGlyGlnValTTrTPProSeRseRcyvlySleAa 189
Db      122  GGGGTGGGCGCCATTCATGCGCAACGCGGAGAAATGGTGGCTCACTCGCGCAAGATCAAC 181
Qy      190  LySeRseRleuProCyAaRPaRAlaProThAlaLeuThPhRThPaRcyvlyAaenLeu 209
Db      182  AAGGCTCTCCCTTGCAAGAGGAGGCTCGAGCGGCTGTGATTCACATTCGACATCGCGCACTG 241
```

```
Qy      210  LVValAlaenlleuLySeRlyAaenlleGlnGlnlleHlellelyeRPhRgLyuseRcy 229
Db      242  AAGTGGAGATCTGAATAATCTGAAACAGCCAAAGATACATGACATGCTGAGGATTC 301
Qy      230  ThraenValAlaAseRleuAaenlleAaenlleAaenlleAaenlleAaenlleAa 249
Db      302  GCAAAATGCCCTGTGGCGGGTGTTCATCAGAGGCGGCTGGGACCGCCCTAAGCTGAC 361
Qy      250  GLyAlHleValSeRanThRlnTyrGlnlleGlnlleSeRseRThRllelleGlyThR 269
Db      362  GGCATTCACATCATCTCGACGAAAGATGTACCGCTGACAGACTGCAAGATCAAGAGG 421
Qy      270  AaRPaRcyvllSeRlleValAseRgLySeRglnAaenValGlnAaenlleThraenlleThRcy 289
Db      422  GACGATGATGATCAATTCAGAAACGGAGCTCAACACTTCATGTTCCAAAGTTAACTG 481
Qy      290  GLyProGlyHleGlylleSeRlleGlySeRleuglySeRgLyAaenSeRglnAlaTyrVal 309
Db      482  GGTCCAGGCGCATGGATTCAGATCGGAGACCTTAGAGAGACAGCAACTCAAGAGCAAGTC 541
Qy      310  SeRanValThRValaenGlnAlaLySlelleGlyAlaGlnAaenGlyValHrglyle 329
Db      542  TCCGCGATCACCATGATTCAGTGCACATCAACGCGCAACCAACGAGCAGCAGCATCAAG 601
Qy      330  ThRTPRglnGlyGlySeRgLyGlnAlaSeRanllelyeRPhRleuAaenValGlnMetGln 349
Db      602  ACGTACGAGGAGGCGGAGATTCGCAAGACATCACTTCGAAATATGTTATGTAC 661
Qy      350  AaRValLyTyRProIlelleleAaRglnAaNTYTCyAaRPaRgValGluProCyv1le 369
Db      662  GACGTCAAAACCCCATATCATCTCCAGACGAACTACTGCGACAGGCTAAGCATCGGA 721
Qy      370  GlnGlnPhSeRAlaValGlnValLyAaenValValTyrGlnAaenllelyseRlyThRse 389
Db      722  GAAACAAGATTCAGAGGTTCAGCGGTGTCTTCAAGAACTTAGAGGAGGAGGAGC 781
Qy      390  AlaThRlyeValAllelyeRheaRPySeSerThraenPhRProCyseGlnGlylle 409
Db      782  AGTACCAAGAGCGCATCAAGATGAATCTGAGTGAAGAGTCCCATGCGCAAGCATTACC 841
Qy      410  MetGlnAaenlleAaenlleuValGlyuseRgLyLyV--ProSeRglnAlaThRcyvly 428
Db      842  TTGCGAAGATGACCTGAAATTCAGACGAGCGGACGCGCAACGAGCAAGAGCACTTTCAG 901
Qy      429  AaenlleAaenlleAgluHleValThRProHle--CyethRseleuglu 446
Db      902  AATGCAAAATGGAAGAAATTTGGAGACATTCGCCCGGACGCTTGACATGCAATAAG 958

RESULT 15
US-10-437-963-52002
/ Sequence 52002, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 52002
/ LENGTH: 1573
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_54340C.1
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 4, 2005, 12:56:14 ; Search time 4970 Seconds  
(without alignment)

4455.540 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390  
Sequence: 1 MVIQRNSILLIITIPASSIS.....VTPHCTSLSEIDELALNY 457

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

MODEL=frame+ p2n\_model -DEV=xlp  
-O=/cgn2.1/USPTO spoool.p/US10691374/runat 28022005 120705 20986/app\_query.fasta\_1.647  
-DB=genemb1 -QFMT=fastcap -SUFFIX=rgc -MINMATCH=0.1 -LOOPT=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdt -LIST=45  
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10691374 @CGN 1 1 5600 @runat 28022005 120705 20986 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGIO  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmb1:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pal:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_pro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	1617	8 LEPGR	X05656 Tomato mRNA
2	2390	100.0	1621	6 A15981	A15981 L.esculentu
3	2390	100.0	1621	6 AX062336	AX062336 Sequence
4	2390	100.0	1621	8 LEPG2AR	X04583 Tomato mRNA

5	2390	100.0	1624	6 I01809	I01809 Sequence 1
6	2390	100.0	1636	6 A24194	A24194 L.esculentu
7	2390	100.0	1636	6 AR364905	AR364905 Sequence
8	1506	63.0	7456	8 LEPOLYGA	X14074 Tomato gene
9	1506	63.0	7456	8 TOMPGANA	M37304 Tomato poly
10	1282	53.6	7456	8 AF152758	AF152758 Actinidia
11	1204.5	50.4	1595	8 AY043233	AY043233 Vitis vin
12	1153	48.2	1725	8 AV0POLYCAL	L06094 Avocado pol
13	1149	48.1	1364	8 AY078936	AY078936 Arabidops
14	1136	47.5	1755	8 PAPOLYGA	X66426 P.americana
15	1130	47.3	1765	8 BNA250919	AJ250919 Brasica
16	1128	47.2	1767	8 AF062467	AF062467 Cucumis m
17	1127	47.2	1296	6 AX412563	AX412563 Sequence
18	1127	47.2	1296	6 AX651724	AX651724 Sequence
19	1127	47.2	1296	6 BT005376	BT005376 Arabidops
20	1127	47.2	1296	6 AK117942	AK117942 Arabidops
21	1115.5	46.7	1280	6 AX651722	AX651722 Sequence
22	1113	46.6	1599	8 BNPALACR	X95800 B.napus mRN
23	1113	46.6	1631	6 A62402	A62402 Sequence 1
24	1106	46.3	1302	8 BRP428543	AJ428543 Brasica
25	1103	46.2	1657	6 A57101	A57101 Sequence 5
26	1103	46.2	1657	6 AX016328	AX016328 Sequence
27	1103	46.2	1657	6 BNSAC66	Z49971 B.napus of
28	1092	45.7	1773	8 MAUP14A	L27743 Malus domes
29	1074	44.9	1795	8 AB084461	AB084461 Pyrus com
30	1061.5	44.4	1208	8 R1M4147	AJ224147 Rubus ida
31	1059	44.3	599	8 CPAS05750	AJ565750 Carica pa
32	1044	43.7	1641	8 PC0504855	AJ504855 Pyrus com
33	1022.5	42.8	1269	6 AX653992	AX653992 Sequence
34	1011	42.3	1182	6 AX653993	AX653993 Sequence
35	1005	42.1	602	8 LBS505947	AJ505947 Lycopersi
36	999	41.8	1344	6 AX654086	AX654086 Sequence
37	925	38.7	1603	8 AF138858	AF138858 Lycopersi
38	914	38.2	1366	8 AY142668	AY142668 Arabidops
39	914	38.2	1471	8 AY046002	AY046002 Oryza sat
40	911	38.1	1450	8 AK105520	AK105520 Oryza sat
41	868.5	36.3	87792	8 AP003837	AP003837 Oryza sat
42	852.5	35.7	4440	8 PPRG	X77231 P.persea P
43	849	35.5	87792	8 AP003837	AP003837 Oryza sat
44	849	35.5	120823	8 AP003748	AP003748 Oryza sat
45	841.5	35.2	1236	6 AX652801	AX652801 Sequence

#### ALIGNMENTS

RESULT 1	LEPGR	1617 bp	mRNA	linear	PLN 12-SEP-1993
LOCUS	Tomato mRNA for polygalacturonase (PG, EC 3.2.1.15).				
DEFINITION	LEPGR				
ACCESSION	X05656				
VERSION	X05656.1	GI:19297			
KEYWORDS	polygalacturonase.				
SOURCE	Lycopersicon esculentum (tomato)				
ORGANISM	Lycopersicon esculentum				
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.				
AUTHORS	Sheehy,R.E., Pearson,J., Brady,C.J. and Hiatt,W.R.				
TITLE	1 (bases 1 to 1617)				
JOURNAL	Mol. Gen. Genet. 208, 30-36 (1987)				
REFERENCE	2 Rogers,H.J., Allen,R.L., Hamilton,W.D. and Lonsdale,D.M.				
AUTHORS	Pollen specific cDNA clones from Zea mays				
TITLE	Biotechm. Biophys. Acta 1089 (3), 411-413 (1991)				
JOURNAL	91316147				
MEDLINE	1859845				
PUBMED	3 (bases 1438 to 1438)				
REFERENCE	Hiatt,W.R.				
AUTHORS	Direct Submission				
TITLE	Submitted (06-OCT-1987)				
JOURNAL	Location/Qualifiers				
FEATURES	1..1617				
SOURCE					

```

/organism="Lycopersicon esculentum"
/db_type="mRNA"
/db_xref="taxon:4081"
/clone="F1"
47..1420
/note="unnamed protein product; polygalacturonase (AA
1-457)"
/codon_start=1
/protein_id="CAA29148.1"
/db_xref="GI:19238"
/db_xref="GOA:P05117"
/db_xref="UniProt/Swiss-Prot:P05117"
/translation="MVIQNSILLLIIIPASSISTGCSNYIDNLFKQYVDNLLRQEP
ADHFOALVSLSKIESNNIDKVDKMGIVINVLSPGAKGDKETVDNLRANMBA
CSRTPVQFVPPKRNKVLKQIIPSGCRSISVYKIFGSLBASKISDTRKRLTAF
DSYONLVYGGGRTINGQYVWPSSCKINSLPCRDAPLALTYWNCNKLVNLSKN
AQQHTIFRSECTNVASNLMIINASAPNTDGVASVNTQYIQISDPIIGGDCISIV
SGSONVATNITCGPHGISIGISGNSSEAVSVNTVAKIIGANGVRLKTMQGG
SGQASNIKPLNVEDVKPPIIDONVCEBPCIOQFSAVQVNVVENEIKGTSATK
VAIKFDCTNPFCEGIIMENINLVSGSGSEATCKNVHNAEHTPHCTSLSESD
EALLVNY"
1438
/note="c was g in [1]"
/citation=[1]
1465..1470
/note="pot.polya signal"
1597..1603
/note="pot.polya signal"

ORIGIN

Alignment Scores:
Pred. No.:      6.24e-192      Length:      1617
Score:          2390.00      Matches:      457
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:    100.00%      Indels:      0
DB:              8          Gaps:          0

US-10-691-374-2 (1-457) x LEPGR (1-1617)

QY      1 MetValIIeGlnAArgAnSerIleLeuLeuLeuIIeIIeIlePheAlaSerSerIleSer 20
Db      47 AAGCTTATCCAAAGAAATGATTTCTCTTCATATTAATTTGCTTCATCAATTTCA 106
QY      21 ThrCysArgSerAsnValIIeAspAspAsnLeuPheGlnValIlyrAspAsnIleLeu 40
Db      107 ACTGTAGAAGCAATGTATTGATGACATTTATCAACAAGCTTATGATATAATTTCTT 166
QY      41 GluGlnGluPheAlaHisAspPheGlnAlaIlyrLeuSerIlyrLeuSerIlyAsnIleGlu 60
Db      167 GAACAAAGAAATTTGCTCATATTTTCAAGCTTATCTTTCTTATTGACAAAATATTGAA 226
QY      61 SerAsnAsnAsnIIeAspIlyValAspIlyAsnGlyIleIlyValIIeAsnValLeuSer 80
Db      227 ACGCACAATATATTTGACAAAGTTGATATAAATGGGATTAAAGTAAATGATCTTAC 286
QY      81 PheGlyAlaIyGlyAspGlyIlySerIlyrAspAsnIleAlaPheGlnGlnAlaIlyrAsn 100
Db      287 TTTGAGGCTAAGGCTGATGAAACATATGATATATTGACATTTGACCAAGCATGAAAT 346
QY      101 GluAlaCysSerSerArgThrProValGlnPheValIleProValAsnIlyAsnIlyLeu 120
Db      347 GAAAGATGTTCACTAGAAACACCTGTTCAATTTGGTGCTCTTAAACAAAGAAATTTATCTT 406
QY      121 LeuIlyGlnIleThrPheSerGlyProCysArgSerSerIleSerValIlyIlePheGly 140
Db      407 CTCAGCAAAATCACCTTTTCAGGTCATGCGAGATTTCTATTTCATTAAGTAAATTTTGGGA 466
QY      141 SerLeuGlnAlaSerSerIlyIleSerAspIlyrIlyAspAspArgIleuIlyIleAlaPhe 160
Db      467 TCTTAAAGAGATCTAGTAAATTTCAAGATTAACAAGATGAAGGCTTTGGATGCTTTT 526
QY      161 AspSerValGlnAsnLeuValIleGlyIlyGlyIlyThrIleAsnGlyAsnGlyGlnVal 180

```

```

Db      527 GATAGTGTTCAAAATTAAGTTGTTGAGAGAGAGAACTATCAATGGCAATGACAGTA 586
QY      181 TTTTTPProSerSerCysIlyIleAsnIlySerIleuProCysArgAspAlaProThrAla 200
Db      587 TGGTGGCAAGTTCTTGCAAAATTAATTAATCACTGACCATGAGGAGGACCAACGGCC 646
QY      201 LeuThrPheTrpAsnCysIlyAsnLeuIlyValIleAsnLeuIlySerIlyAsnIleGln 220
Db      647 TTAACCTTCTGGAATTTGCAAAATTTGAAGATGATATCTAAAGATTAATACACAA 706
QY      221 GlnIleHisIleIlyPheGlyIlySerIlyrAsnValIleAlaSerAsnLeuIleAsn 240
Db      707 CAATTCATATCAAAATTTGAGTCAAGCACTATATGTTGAGCTTCAAAATTTGATGATCAT 766
QY      241 AlaSerAlaIySerProAsnThrAspIlyValHisAlaSerAsnThrGlnIlyIleGln 260
Db      767 GCTTCAGCAAGAGGCCAAATTACTGATGAGTCCATGATCAATCAATATATTTCA 826
QY      261 IlySerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValIlySerGln 280
Db      827 ATATCTGATACTATATTGAAACAGTGAATGATGATTTCAATTTGTTCTGATCTCAA 886
QY      281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300
Db      887 AATGTGACGGCCAAATATTAATCTTGCTCAGTCAAGTATGATTAAGTATTTGAAAGCTTA 946
QY      301 GlySerGlyAsnSerGlnAlaIyIyValSerAsnValThrValAsnGlnAlaIyIle 320
Db      947 GGAATCGAATTTCAAGAAAGCTTATGTTGTTATCTTAATCTTAATTAAGCAAAATTTATC 1006
QY      321 GlyValGlnAsnGlyValAlaGlyIleIlyThrIlyGlnGlyIlySerGlyGlnAlaSerAsn 340
Db      1007 GGTGCGCAAAATTTGAGTGAATCAAGCTTGGCAGGAGATCTGGAACAAGTACGCAAC 1066
QY      341 IlyIlyPheLeuAsnValGlnMetGlnAspValIlyrProIleIleIleAspGlnAsn 360
Db      1067 ATCAATTTCTGATGTGGAATGCAAGATCAAGCTTAAAGTATTAATTAATTAAGCAAAATC 1126
QY      361 TyrCysAspArgValGlnProCysIleGlnGlnPheSerAlaValGlnValIlyAsnVal 380
Db      1127 TATGTATGATGAGTTGAACCATGATATCAACAAGTTTCAACAAGTCAAGTGAAGAAATGTG 1186
QY      381 ValIlyrGlnAsnIleIyGlyThrSerAlaThrIlyValAlaIleIlyPheAspCysSer 400
Db      1187 GTGTATGAAATATCAAGGCAAGGCAAGTCAACAAGGCTGCCATTAATTTGATTTGCAGC 1246
QY      401 ThrAsnPheProCysGlnGlyIleIleMetGlnAsnIleAsnLeuValIlyGlyIlySerGly 420
Db      1247 ACAACCTTTCATGTGAAGAAATTAATATGAGAAATTAATTAATTTAGAGGGAAGTGA 1306
QY      421 LysProSerGlnAlaThrCysIlyAsnValHisPheAsnAsnAlaGlnHisValIlyrPro 440
Db      1307 AAACCATCAAGAGCTACGTGCAAAATTTGTCATTTTAAACAATGCTCAACAATTTTACACA 1366
QY      441 HisCysThrSerLeuGlnIlySerGlyAspGlnAlaLeuLeuIlyrAsnIlyr 457
Db      1367 CACTGCACTTCACTAAGAAATTTCAAGAGATGAAGCTTTTGTATATATTAAT 1417

RESULT 2
A15981      1621 bp      RNA      linear      PART 02-MAR-1994
LOCUS      Lycopersicon esculentum
DEFINITION  Lycopersicon esculentum
ACCESSION  A15981
VERSION    A15981.1 GI:490039
KEYWORDS   Lycopersicon esculentum (tomato)
ORGANISM   Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE  1 (bases 1 to 1621)
            Bridges,I.G., Schuch,W.W. and Grierson,D.
AUTHORS

```

TITLE Anti-sense regulation of plant gene expression  
JOURNAL Patent: EP 0271988-A 1 22-JUN-1988;  
IMPERIAL CHEMICAL INDUSTRIES PLC  
FEATURES Location/Qualifiers

## source

1..1621  
/organism="Lycopersicon esculentum"  
/mol\_type="unassigned RNA"  
/strains="Alisa Craig"  
/db\_xref="taxon:4081"  
/clone="PTOM 6"  
/cissue\_type="fruit"  
/dev\_stage="fruit ripe"  
51..1424  
/codon\_start=1  
/product="polygalacturonase-2a"  
/protein\_id="CAA01256.1"  
/db\_xref="GI:490040"  
/db\_xref="GOA:P05117"  
/db\_xref="UniProt/Swiss-Prot:P05117"  
/translation="WVIGRNSILILITIPASSISTCRSNVIDDNLFKOYDNIIEQEF  
AHDFOAYLSYKSNIESNNIDKVDKNGIKVINLVSFGAKGDKTYDNIIEQANEA  
CSRTVPQVVPVKNNYLLKQITFSGPKRSISVRI FGLSEASXISDYKORLWIAF  
DSVOLVVGGGTNGVQWSPSSCKINKSLPCRDAPLALFMCKLKYNNLSKN  
AQOIHKEPSCNVAASNMJNASKSPNIDGVHNTQYIOISDTIIGTGDICISIV  
SGSNVQNTNITCGPHGISTGSGNSSEAVSVTVNNAKIIIGAEVGRKTMGG  
SGQSNITKFLAVEMQDVKPIITIDNYCDRPEPCIQPSAVQKRVYENIKGTSATK  
VAIKPDCSTNPFCEGIIMENINLVGESKPEATCKNVPNNAEHVTCHTSLISED  
BALNLYN"

## CDS

mat\_peptide  
264..1421  
/product="polygalacturonase-2a"  
615..623  
/note="pot. N-glycosylation site"  
misc\_feature  
768..776  
/note="pot. N-glycosylation site"  
misc\_feature  
906..914  
/note="pot. N-glycosylation site"  
misc\_feature  
981..989  
/note="pot. N-glycosylation site"  
misc\_feature  
1598..1606  
/note="pot. N-glycosylation site"  
misc\_feature  
1598..1606  
/note="pot. polyA signal"  
polyA\_site  
1621  
/note="polyA site"

## ORIGIN

## Alignment Scores:

Pred. No.: 6,266-192 Length: 1621  
Score: 2390.00 Matches: 457  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-691-374-2 (1-457) x AL5981 (1-1621)

QY 1 MetValIIEGlnARgsenSerIleleuLeuIleIleIlePheAlaSerSerIleSer 20  
DB 51 ATGGATTATCCAAAGAAATGATATCTCTTCATTATATATTTTGCTTCATCAATTTCA 110  
QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheValIleValIleAspAsnIleLeu 40  
DB 111 ACTTGTAGAGCAATGTTATTTGATGACATTTTATTCAAACAAGTTTATGATAATTTCTT 170  
QY 41 GIUGInGluPheAlaHisPhePheGlnAlaIleLeuSerIleLeuSerIleLeuSerIleGlu 60  
DB 171 GAACAAGAAATTTGCTCATGATTTTTCACGCTTATCTTTCTTATTTGACAAAAATATTGAA 230  
QY 61 SerAsnAsnAsnIleAspIleValIleAspIleAsnGlyIleIleValIleAsnValIleAsnValIleuSer 80  
DB 231 ACCAACAAATATATATGACAAAGGTGATTAATAATGGGATTAAGTGATTAATGACTTAC 290  
QY 81 PheGluAlaIleGluIleAspGluIleAspGluIleAspAsnIleAlaPheGluGlnAlaTrpAsn 100  
DB 291 TTTGGAGCTAAGGCGATGGAABAAACATATGATTAATTTGCAATTTGACCAAGCATGGAAAT 350

QY 101 GIUAlaCysSerSerArgThrProValGlnPheValValProIleAsnIleAsnIleLeu 120  
DB 351 GAAGCATTTTCATCTAGAACACCTGTTCAATTTTGTTCTTCAAAAACAAAGATATATCTT 410  
QY 121 LeuIleGlnIleLeuThrPheSerGluProCysArgSerSerIleSerValIlePheGlu 140  
DB 411 CTCAGCAAAATCACCTTTTTCAGTTCATGCAATCTTCTATTTCACTAAGATTTTGTGA 470  
QY 141 SerLeuGluAlaSerSerIleSerIleSerIleValIleAspArgArgLeuTrpIleAlaPhe 160  
DB 471 TCCTTAGAAGCATCTAGTAATTTTCAAGATCAAAAGATAGAGGCTTTGGATTCCTTTT 530  
QY 161 AspSerValGlnAsnLeuValAlaGluIleGluIleIleAsnGluIleAsnGluIleVal 180  
DB 531 GATAGTGTTCAAATTTTGTGTGGAGAGAGGACCTATCAATGGCAATGACAAAGTA 590  
QY 181 TrpTrpProSerSerCysIleIleAsnIleSerLeuProCysArgAspAlaProThrAla 200  
DB 591 TGGTGGCCAAAGTCTTTCGAAATTAATAATCATCTGCCATGCAAGGATGCAACCAAGGCC 650  
QY 201 LeuThrPheTrpAsnCysIleAsnLeuIleValAsnAsnLeuIleSerIleAsnAlaGln 220  
DB 651 TTAACCTTCTGGAAATTCGAAAAATTTGAAGTGAATTAATCTAAGAGTAAAGTAAATGCACA 710  
QY 221 GlnIleHisIleIlePhePheGluSerCysThrAsnValValAlaSerAsnLeuMetIleAsn 240  
DB 711 CAATTCATATCAAAATTTGAGTCAAGCACTAATGTTGAGCTCAAAATTTGATGATCAT 770  
QY 241 AlaSerAlaIleSerProAsnThrAspGluValIleHisValSerAsnThrGlnIleGln 260  
DB 771 GCTTACGCAAGAGGCCAAATATCTGATGAGTCCATGATCAAAATCTCAATATTTCA 830  
QY 261 IleSerAspThrIleIleGluIleThrGluAspAspCysIleSerIleValIleSerGln 280  
DB 831 ATATCTGATCTAATATTGGAACAGATGATGATTTGATTTCAATTTGTTCTGATCTCA 890  
QY 281 AsnValGlnAlaThrAsnIleThrCysGluIleProGluIleGluIleSerIleGluSerLeu 300  
DB 891 AATGTGCGGCCCAAAATTTTACTTGTGTCCAGGTCAAGTCAATGATTAAGTATGGAAGCTTA 950  
QY 301 GlySerGluAsnSerGluAlaIleValIleSerAsnValIleHisValAsnGluAlaIleIle 320  
DB 951 GGATCTGCAAAATTCGAAAGCTTATGCTATATGTTACTGTAATGAAAGCCAAATATATC 1010  
QY 321 GluAlaGluAsnGluValArgIleIleThrTrpGlnIleGlySerGluAlaSerAsn 340  
DB 1011 GGTCGCAAAATGAGATGATCAAGACTTGCGCAGGAGGATCTGCAACAACTAGCAAC 1070  
QY 341 IleIlePheLeuAsnValGluMetGlnAspValIleIleTrpProIleIleIleAspGlnAsn 360  
DB 1071 ATCAAAATTTCTGAATGTGAAATGCAAGCGTTAAGTATCCATATATTATTAACCAAAAC 1130  
QY 361 TyrCysAspArgValGluProCysIleGlnIlePheSerAlaValGlnValIleAsnVal 380  
DB 1131 TATTTGATCGAGTTGAACCATGATATCAACAGTTTTCAGCAAGTTCAAGTGAATATG 1190  
QY 381 ValIleGluAsnIleIleGluIleThrSerAlaThrIleValAlaIleIlePheAspCysSer 400  
DB 1191 GTGTATGAAATATCAAGGCGCAAGTCAACAAAGGCGCCATTAATTTGATTGACAC 1250  
QY 401 ThrAsnProCysGluGluIleIleIleMetGluAsnIleAsnLeuValGluIleSerGly 420  
DB 1251 ACAAACTTTCATGGAAGGAATTTAATGGAATTAATAATTTGATGAGGGAAGTGA 1310  
QY 421 LysProSerGluAlaThrCysIleAsnValIlePheAsnAlaGluIleValIleThrPro 440  
DB 1311 AAACCATCAAGGCTACGTCGCAAAATGCTATTTTCAAAAGCTGAACATTTACACA 1370  
QY 441 HisCysThrSerLeuGluIleSerGluAspGluAlaLeuLeuIleIleIleIleIleIle 457  
DB 1371 CACTGCACTTCACTAGAAATTTCAAGATGATGAAGCTTTTGTATATATAT 1421

```

RESULT 3
AX062336      AX062336      1621 bp      DNA      Linear      PAT 24-JAN-2001
LOCUS         Sequence 1 from Patent WO0078982.
DEFINITION    AX062336
ACCESSION     AX062336
VERSION       AX062336.1  GI:12540229
KEYWORDS
SOURCE        synthetic construct
ORGANISM      other sequences; artificial sequences.
REFERENCE     1
AUTHORS       Christensen,T.M. and Kreiberg,J.D.
TITLE         Process for the enzymatic modification of pectin
JOURNAL       Patent: WO 0078982-A 1 28-DEC-2000;
              DANISCO A/S (DK)
FEATURES
  source       1..1621
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="cDNA clone"
ORIGIN
Alignment Scores:
Pred. No.:    6.26e-192      Length:    1621
Score:        2390.00       Matches:    457
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:   100.00%       Indels:    0
DB:            Gaps:        0
US-10-691-374-2 (1-457) x AX062336 (1-1621)
QY      1 MetValIIIGLhAryAsnSerIleLeuLeuIleIleIlePheAlaSerSerIleSer 20
DB      51 ATGGTATTCAGAAAGGATGGATATTCCTTCATTAATTAATTTGCTTCATCAATTTCA 110
QY      21 ThrCyArgSerAsnValIleAspAspAsnLeuPheIleValIleAspAsnIleLeu 40
DB      111 ACTTGAGAGCAATGTATTGATGACATTTATTCAAACAGTTTATGATATATTTCTT 170
QY      41 GIuGIuGIuPheAlaHisAspPheGlnAlaIleLeuSerIleLeuSerIleLeu 60
DB      171 GAACAAAGAAATTTGCTCATATATTTCAAGCTTATCTTTCTTATTTGAGCAAAATTTTAA 230
QY      61 SerAsnAsnAsnIleAspIleValAspIleAspIleLeuValIleAsnValIleuSer 80
DB      231 AGCAACAAATATATTCAGCAAGGTTGATTAATAATGGATTAAAGTATTAATCTTAC 290
QY      81 PheGlyAlaIleGlyIleAspGlyIleThrIleAspAsnIleAlaPheGIuGIuAlaIleProAsn 100
DB      291 TTTGAGGCTTAAGGCTGATGAGAAACATATGATTAATGACATTGAGCAAGCATGAAAT 350
QY      101 GIuAlaCySerSerArgThrProValGlnPheValIleProValAsnIleAsnIleu 120
DB      351 GAAGCATGTTCTATTCAGAACACCTGTTCATTAATTTGGTTCCTTAATAACAGAAATATCTT 410
QY      121 LeuIleGlnIleThrPheSerGlyProCyArgSerSerIleSerValIleIlePheGly 140
DB      411 CTCAGCAAAATCACCTTTTCAGGTCAGTCAGCATCTTTCTATTTCAGTAAAGATTTTGGCA 470
QY      141 SerLeuGIuAlaSerSerIleIleSerAspIleIleAspIleValIleAspIleValIleAlaPhe 160
DB      471 TCTTAAAGACATCTAGTAATAATTTTCAGACTTAACAAAGATAGAGGCTTTGGATTGCTTTT 530
QY      161 AspSerValGlnAsnLeuValIleGlyIleGlyIleThrIleAsnIleAsnIleGlyValIle 180
DB      531 GATAGTGTTCAAAATTTAGTCTTGAGAGAGAGAGACCTATCAATGGCAATGGACAAAGTA 590
QY      181 TrpTrpProSerSerCyAlaIleAsnIleSerIleuProCyArgAspAlaProThrAla 200
DB      591 TGGTGGCCAAAGTTCTTGCAAAATTAATTAATCACTGCATGCAAGGATGCAACGAGCC 650

```

```

QY      201 LeuThrPheTrpAsnCyAlaIleAsnLeuIleValIleAsnIleLeuSerIleAsnIleGln 220
DB      651 TTAACTCTTGGAATTCGAAAAATTTGAAGGATTAATCTTAAGAGTAAAAATGCACAA 710
QY      221 GlnIleHisIleIlePhePheGlyIleSerIleAsnValIleAlaSerAsnIleuMetIleAsn 240
DB      711 CAATTCATATCAAAATTTGAGTCAGCACTATATGTTGAGCTTCAAAATTTGATGATCAAT 770
QY      241 AlaSerAlaIleSerProAsnThrAspGlyValIleHisIleSerAsnThrGlnTyrIleGln 260
DB      771 GCTTCAGCAAAAGGCCAAATATCTGATGAGATGCCATGATCAAAATCTCAATATATTTCA 830
QY      261 IleSerAspThrIleIleGlyThrGlyAspAspCyIleSerIleValIleSerGlySerGln 280
DB      831 AATATCTGATCTATATATGGAACAGTGATATGATATTCATTAATTTGATTCGATCTCA 890
QY      281 AsnValGlnAlaThrAsnIleThrCyGlyIleProGlyIleGlyIleSerIleGlySerIleu 300
DB      891 AATGTCAGGCCCAAAATATTAATCTGTGTCCAGGTCATGGTATTAATGAAAGGCTTA 950
QY      301 GlySerGlyAsnSerGlyAlaIleValSerAsnValThrValAsnGlyAlaIleIle 320
DB      951 GGATCTGGAATTTGGAAGCTTATATGTGTCTAATGTTACTGTAATGAAGCCAAATATATC 1010
QY      321 GlyAlaGlyAsnGlyValArgIleIleIleThrIleGlnGlyIleSerGlyGlnIleSerAsn 340
DB      1011 GGTGCCGAAATTTGAGATGATCAAGCTTGCGCAGGAGGATCTGACAAAGCTACGCAAC 1070
QY      341 IleIlePheLeuAsnValIleGlyIleAspValIleIleIleIleIleIleIleIleIleIle 360
DB      1071 ATCAAAATTTTGAATGTGGAATTCGAAATTCAGACCTTAAGTATCCATTAATTAATGACCA 1130
QY      361 TyrCyAspArgValIleProCyAlaIleGlnIlePheSerAlaValGlnValIleAsnVal 380
DB      1131 TATGTGATCGAAGTTGAACCATGATATCAACAGTTTCACAGTTCAAGGAAATATGTG 1190
QY      381 ValTyrGluAsnIleIleGlyIleThrSerAlaThrIleIleValIleIlePheAspCySer 400
DB      1191 GTGTATGAAATATCAAGGACAGTCACAAAGGTCGCATTAATAATTTGATGAGC 1250
QY      401 ThrAsnPheProCyGlyGlyIleIleMetGluAsnIleAsnLeuValIleGlyIleSerGly 420
DB      1251 ACAACCTTTTCATGATGAAGGATTAATATGAGAAATTAATTAATTAAGGAGGAAAGTGA 1310
QY      421 LysProSerGlyAlaThrCyAlaIleAsnValIlePheAsnAsnIleGluIleValIleThrPro 440
DB      1311 AATCCATCAAGGCTAGTCGCAAAATGTCATTTTAACAATGCTGAAACATGTTACACCA 1370
QY      441 HisCyThrSerLeuGluIleSerGlyAspGlyAlaIleLeuIleIleIleIleIleIleIle 457
DB      1371 CACTGCACCTTCATTAAGAAATTTTCAGAGATGAAGCTCTTTTGTATTAATTAAT 1421

```

RESULT 4  
 LEPG2AR 1621 bp mRNA Linear PLN 09-SEP-2004  
 LOCUS Tomato mRNA for polygalacturonase-2a (PG-2a).  
 DEFINITION X04583  
 ACCESSION X04583  
 VERSION X04583.1 GI:19291  
 KEYWORDS glycoprotein; polygalacturonase; polygalacturonase-2a; signal peptide.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 1621)  
 Grierson,D., Tucker,G.A., Keen,J., Ray,J., Bird,C.R. and Schuch,W.  
 Sequencing and identification of a cDNA clone for tomato polygalacturonase  
 Nucleic Acids Res. 14 (21), 8595-8603 (1986)  
 JOURNAL MEDLINE  
 MEDLINE 87066731  
 PUBMED 3786135  
 COMMENT The predicted N-terminal AA sequence of PG-2a occurs at AA72. Since

the ATG at bp 51 codes for the only inframe methionine preceding the N-terminal sequence of PG-2a, it is probable that the mRNA is translated with a 71aa pre-sequence that is subsequently cleaved to give a 41,828D mature protein.

Data kindly reviewed (22-APR-1987) by D. Grierson.

## FEATURES

## SOURCE

1..1621

Location/Qualifiers

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/strain="Alisa Craig"

/cultivar="Alisa Craig"

/db\_xref="taxon:4081"

/clone="PTOM 6"

/tissue\_type="fruit"

/dev\_stage="fruit ripe"

51..1424

/note="unnamed protein product; precursor polypeptide (AA

-71 to 386)"

/codon\_start=1

/protein\_id="CAA28254.1"

/db\_xref="GI:19292"

/db\_xref="GOL:P05117"

/translation="WVIOKNSILILIIIFASISIFCRSNVIDNLFKQVNDILEOE

AHPQAVLSYLSKINIESNNIDKDKNGIKVINLSPAGKGGKTYDNIAPQAWNEA

CSSTRPVQPVYPRKKNYILKQTFSGPCRSISVYKIFSLSSKISDYKORRLIAR

DSQNLVVGGGTINNGQVWMPSSCKINKSLPCBDAFTALIFMCKNLKNNLSKN

AQOIHFKESCTGNVNASNLINAKSPNTDGVHSNQTQISDTIIGTGDICISV

SGSNVQANITTCGPHGISIGLSGNSSEAVSVTVNEAKIIGAENGVRKTKQGG

SGQASNIKFLVEMQDVKPIIIDQNYCDRVPICQPSAVQKVVENIKGTSAIK

VAIKEDGNTNPFCEGIIMENINLVGESKPSBATKQNFNNVAHVTHCHTSLEISD

EALLYNY"

264..1421

/product="mature polygalacturonase (AA 1-386)"

51..263

/note="put. pre-peptide (AA -71 to -1)"

615..623

/note="pot. N-glycosylation site"

768..776

/note="pot. N-glycosylation site"

906..914

/note="pot. N-glycosylation site"

981..989

/note="pot. N-glycosylation site"

1598..1606

/note="pot. polyA signal"

1621

/note="polyA site"

## ORIGIN

## Alignment Scores:

Pred. No.: 6.26e-192 Length: 1621  
Score: 2390.00 Matches: 457  
Percent Similarity: 2390.00 Conservative: 0  
Best Local Similarity: 100.00 Mismatches: 0  
Query Match: 100.00 Indels: 0  
DB: 8 Gaps: 0

US-10-691-374-2 (1-457) x LEPG2AR (1-1621)

QY 1 MetValIleGlnArgAsnSerIleLeuLeuLeuIleIlePheAlaSerSerIleSer 20  
DB 51 ATGGTTATTCAGAAAGAAATGATATCTCCCTTCATTAATTAATTTGCTTCATCAATTTCA 110  
QY 21 ThrCyAspSerAsnValIleAspAspAsnLeuPheValIleTyrAspAsnIleLeu 40  
DB 111 ACTTGTAAGACAAAGTATTGATGCAATTAATTCAGCAAGTTATATATAATTTCTT 170  
QY 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerIleAsnIleGlu 60  
DB 171 GAACAAGAAATTTGCTTCATGATTTTCAAGCTTATCTTTCTTATTGAGCAAAAATATTGAA 230  
QY 61 SerAsnAsnAsnIleAspIleValAspIleAsnGlyIleIleValIleAsnValLeuSer 80

DB 221 AGCAACATATATATGACAGCTTGATATAAATGGATTTAAAGTATTATGACTTTAGC 290  
QY 81 PheGlyAlaIleGlyAspGlyIleYserThrTyrAspAsnIleAlaPheGlnAlaTyrAsn 100  
DB 291 TTTGGAGCTTAAGGGATGAGTGAATAAATCATATGATTAATATTGCAATTTGAGCAACATGAAAT 350  
QY 101 GluAlaCySerSerSerArgThrProValGlnPheValProValAsnIleYserIleLeu 120  
DB 351 GAAAGCATTTTCATTCAGAACACCTGTTCAATTTGGTCTTCAAAAACAGAAATTAATCTT 410  
QY 121 LeuIleYsgIleIleThrPheSerGlyProCyAspArgSerIleSerValIlePheGly 140  
DB 411 CTCAGCAAAATCACTTTTCAGGTTCAGTCAAGCAATCTTCTATTTCAGTAAAGATTTTGA 470  
QY 141 SerLeuGluAlaSerSerIleYserAspTyrIleYserAspArgArgLeuTrpIleAlaPhe 160  
DB 471 TCCTTAGAAGCATTAATTAATTTCAAGACTTACAAAGATAGAGGCTTTGGATTCCTTTT 530  
QY 161 AspSerValGlnAsnLeuValValGlyGlyIleThrIleAsnGlyAsnGlyGlnVal 180  
DB 531 GATAGTGTTCAAAATTTAGTGTGGAGAGAGAGAACTATCAATGCGCAATGACAAAGTA 590  
QY 181 TrpTrpProSerSerCyAlaIleIleAsnIleYserLeuProCyAspArgAspAlaProThrAla 200  
DB 591 TGGTGGCCAAAGTCTTTCGAAAATTAATTAATTAATCTGCCATGACGAGGATGCCAACAGGCC 650  
QY 201 LeuThrPheTrpAsnCyAlaIleAsnLeuIleValAsnAsnLeuIleYserIleYserIleVal 220  
DB 651 TTAACCTTCTGAAATTTGCAAAAATTTGAAAGGAAATTAATCTTAATAAGTAAATGACACA 710  
QY 221 GlnIleHisIleIleYsgPheGluSerCysThrAsnValValAlaSerAsnLeuSerIleAsn 240  
DB 711 CAATTCATATCAAAATTTGAGTCATGACTTAATGTTGAGCTTCAAAATTTGATGATCAAT 770  
QY 241 AlaSerAlaIleYserProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260  
DB 771 GCTTAGCAAGAAAGCCAAATATCTGATGAGTCCATGATCAATAATCAATATATTCGA 830  
QY 261 IleSerAspThrIleIleGlyThrGlyAspAspCyAlaSerIleValSerGlySerGln 280  
DB 831 ATATCTGATTAATTAATTTGAAACAGATGATGATTTGATTTCAATTTGTTCTGATCTCA 890  
QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300  
DB 891 AATGTGACGGCCACAAATTAATTAATTTGATGATGATGATGATTAATTAATTAATTAATTA 950  
QY 301 GlySerGlyAsnSerGlyAlaTyrValSerAsnValThrValAsnGlyAlaIleYserIle 320  
DB 951 GGATCTGAAATTTCAAGAGCTTAATGATGATTAATGATTAATGATTAATGATTAATGATTA 1010  
QY 321 GlyAlaGluAsnGlyValArgIleYserThrTrpGlnGlyGlySerGlyGlnIleSerAsn 340  
DB 1011 GGTCGCAAAATGAGATTTAGATCAAGACTTGGCAGGAGATCTGACAACTAGCAAC 1070  
QY 341 IleIleYsgPheLeuAsnValGluMetGlnAspValIleYTrpProIleIleIleAspGlnAsn 360  
DB 1071 ATCAAAATTTCTGAATGTGAATGCAAGACGTTAAGTATCCATTAATTAATTAATTAATTA 1130  
QY 361 TyrCyAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValIleYserAsnVal 380  
DB 1131 TATTTGATCGAGTTGAAACCATGATTAATCAACGTTTTCAGCAAGTTCAAGTGAATAATG 1190  
QY 381 ValTyrGluAsnIleYsgIleYserAlaThrIleYsgAlaIleYsgPheAspCySer 400  
DB 1191 GGTATGAGAAATATCAAGGCGCAAGTGCACAAAGGTGGCCATTAATTTGATTCAGC 1250  
QY 401 ThrAsnPheProCysGluGlyIleIleMetGluAsnIleAsnLeuValGlyGluSerGly 420  
DB 1251 ACAAACTTTCATGGAAGGAATTAATTAATGGAATTAATTAATTTAGTGGGGAAGGTGA 1310  
QY 421 LysProSerGluAlaThrCysIleYsgAsnValHisPheAsnAsnAlaGluHisIleValThrPro 440

Db 1311 AACCATCAGAGCTACGTGCAAAAATGTCATTTTAACAAATGCTGAACATGTTACCA 1370

Qy 441 HisCythrSerLeuGluIleSerGluAspGluAlaLeuLeuTyrAsnTyr 457

Db 1371 CACTGACCTTCACTAGAAATTTGAGAGATGAAAGCTCTTTTGTATATATAT 1421

RESULT 5

LOCUS 101809 101809 1624 bp ss-DNA linear PAT 21-MAY-1993

DEFINITION Sequence 1 from Patent US 4801540.

ACCESSION 101809

VERSION 101809.1 GI:269802

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1624)

AUTHORS Hatt, W.R., Sheehy, R.E., Shewmaker, C.K., Kridl, J.C. and Knauf, V.

TITLE Pg 9 gene and its use in plants

JOURNAL Patent: US 4801540-A 1 31-JAN-1989; Calgene, Inc.; Davis, CA

FEATURES

source 1..1624 Location/Qualifiers

/organism="unknown"

/mol\_type="unassigned DNA"

Alignment Scores:

Pred. No.: 6 27e-192 Length: 1624

Score: 2350.00 Matches: 457

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

Gaps: 0

US-10-691-374-2 (1-457) x 101809 (1-1624)

Qy 1 MetValIleGlnArgAsnSerIleLeuLeuIleIlePheAlaSerSerIleSer 20

Db 47 ATGGTATCCAAAGAGATGATATCTCTTCATCATTAATTTTGCTTCATCAATTTCA 106

Qy 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheLeuGlnValTyrAspAsnIleLeu 40

Db 107 ACTGTGAGAGCAAGTATGATATGATGACATTTATTCACAAAGTTATGATATATCTT 166

Qy 41 GlnGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerIleAsnIleGlu 60

Db 167 GAACAAGATTTGCTCATGATTTTCAAGCTTATCTTTCTATTGAGCAAAAATATGAA 226

Qy 61 SerAsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSer 80

Db 227 ACACACAAATATATGACAAAGGTGATATAAAATGGAATTAAGTATTAATGACTTAC 286

Qy 81 PheGlyAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGluGlnAlaTyrAsn 100

Db 287 TTTGAGAGCTAAGGGATGAGAAAACATATGATATATGATTTGAGCAAGCATGGAAT 346

Qy 101 GlnAlaCysSerSerArgThrProValGlnPheValProLysAsnLysAsnTyrLeu 120

Db 347 GAAGCATGTTCACTAGAACCTGTTCAATTTGGTCTCTAAAAACAAGATTTATCTT 406

Qy 121 LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly 140

Db 407 CTCAGCAAAATCACCTTTCAAGTCCATGACAGATCTTCTATTCAGTAAAGATTTTGA 466

Qy 141 SerLeuGlnAlaSerSerLysIleSerAspTyrLysAspArgLysLeuTyrIleAlaPhe 160

Db 467 TCTCTTAGAGACATCTAGAAATTTTCAACATCAAGATGAAAGCTTTGATGCTTTT 526

Qy 161 AspSerValGlnAsnLeuValGlyGlyGlyThrIleAsnGlyAsnGlyGlnVal 180

Db 527 GATAGTGTCAAAATTTAGTTGTTGGAGGAGGAAGAACTATCATGCAATGCAACAAGTA 586

Qy 181 TrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaProThrAla 200

Db 587 TGTGGCCAAAGTTCTTGCAAAATTAATTAATCATCGCCATGACGAGGATGCACCAAGGCC 646

Qy 201 LeuThrPheThrAsnCysLysAsnLeuLysValAsnAsnLeuLysSerIleAsnAlaGln 220

Db 647 TTAACTCTTGGAATTTGCAAAAATTTGAAAGGAAATTAATCTAAAGATTAATAATGCACA 706

Qy 221 GlnIleHisIleLysPheGluSerCysThrAsnValAlaLysSerAsnLeuMetIleAsn 240

Db 707 CAATTCATCATCAAAATTTGAGTCATGCACTAATGTATGAGCTTCAAAATTTGATGATCAT 766

Qy 241 AlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260

Db 767 GCTTACAGCAAGAGCCCAAAATCATCATGATGAGTCCATGATCAAAATCAATCAATATTTCA 826

Qy 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280

Db 827 ATATCTGATCATTAATTTGAGAACAGTGAATGATTTTCAAAATTTTCTGAGATCTCA 886

Qy 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300

Db 887 AATGTCAGAGCCCAAAATATTAATCTTGTCAGGTCAATGATATGATGAAAGCTTA 946

Qy 301 GlySerGlyAsnSerGluAlaTyrValSerAsnValThrValAsnGluAlaLysIleIle 320

Db 947 GGATCTGGAATTTCAAGAGCTTATGTCTAATGTCTAATTTCAATTAATTAATGAAAGCAAAATATTC 1006

Qy 321 GlyAlaGluAsnGlyValArgIleLysThrTyrGlnGlyGlySerGlyGlnAlaSerAsn 340

Db 1007 GGTGCGCAAAATGAGATTTAGATCAAAAGCTTGCGAGGAGGATTTGACAAAGCTTAACAAC 1066

Qy 341 IleLysPheLeuAsnValGluMetGlnAspValLysTyrProIleIleIleAspGlnAsn 360

Db 1067 ATCAAAATTTCTGAATGAGGAATGCAAGATGCAAGTAAAGATCCATTAATTAATGACCAAAAC 1126

Qy 361 TyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValLysAsnVal 380

Db 1127 TATTTGATCGAGTTGAAACCAATGATACAAAGTTTCCACAGCTTCAAGGAAAAATGTG 1186

Qy 381 ValTyrGluAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAspCysSer 400

Db 1187 GTGATGAGAAATATCAAGGCGACAAGTGCACAAAGGTGCAATTAATTTGATGAGC 1246

Qy 401 ThrAsnPheProCysGluGlyIleIleMetGluAsnIleAsnLeuValGlyGluSerGly 420

Db 1247 ACAAACTTTCCATGTGAAGGAATTAATGAGAAATTAATTAATTAATTAATTAATTAATTAAT 1306

Qy 421 LysProSerGluAlaThrCysLysAsnValHisPheAsnAsnAlaGluHisValThrPro 440

Db 1307 AACCATCAGAGGCTACGTGCAAAAATGTCATTTTAACATCTGAAACATGTTACCA 1366

Qy 441 HisCythrSerLeuGluIleSerGluAspGluAlaLeuLeuTyrAsnTyr 457

Db 1367 CACTGACCTTCACTAGAAATTTTCAAGAGATGAAAGCTCTTTTGTATATATAT 1417

RESULT 6

LOCUS A24194 1636 bp DNA linear PAT 04-APR-1995

DEFINITION L.esculentum polygalacturonase clone pTOM6.

ACCESSION A24194

VERSION A24194.1 GI:904404

KEYWORDS

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

REFERENCE 1 (bases 1 to 1636)

AUTHORS Bridges, I.G., Grieson, D. and Schuch, W.W.

TITLE Anti-sense regulation of plant gene expression

JOURNAL Patent: EP 0532060-A 1 17-MAR-1993; IMPERIAL CHEMICAL INDUSTRIES PLC; ZENECA LIMITED

FEATURES  
source

CDS

Location/Qualifiers  
1..1636  
/organism="Lycopersicon esculentum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4081"  
51..1424  
/codon\_start=1  
/product="polygalacturonase"  
/protein\_id="CAA01720.1"  
/db\_xref="GI:904405"  
/db\_xref="GOA:P05117"  
/db\_xref="UniProt/Swiss-Prot:P05117"  
/translation="WVIOHNSILLIITIPASSISITCRSNVVIDNLPKOVYNNILEOF  
AHDPRAYISLAKINIESNNNDKVDKNGIKYINYSFQKSGKGYNDINAPQANNEA  
CSSRTVPVPEKKNYILKQITFSGPCRSSISVKI FQSLSESSKISDYKDRRLIAF  
DSVNLVVGGGGTINGQWVWSSCKINSLPCPDAPALTPMKLKNVNLKSN  
AQOIHKEFSCITNVVASNLIMINAKSPATDVSNTQYIOISPTIIGTGDCISIV  
SGSONVATNITCGPHGISIGLSGNSSEAVSNVTNBAKII GAENGVRKTMQGG  
SGOANIKELINEMQDVKPIIIDQNYCDVRPCTIQPSAVQKVVNTIKGTSATK  
VAIKRDCSTNPFCEGIIMENINLVESGKPEATCKNVHFNNAEHVTHCTSLISED  
EALTYNY"

ORIGIN

Alignment Scores:

Prod. No.: 6 336-192 Length: 1636  
Score: 2390.00 Matches: 457  
Percent Simlarity: 100.00% Conservative: 0  
Best Local Simlarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-691-374-2 (1-457) x A24194 (1-1636)

Qy 1 MetValIIEGInARgAnSerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSer 20  
Db 51 ATGCTTATCCAAAGGATAGTATTCCTTCATATTAATTTTGGCTTCATCAATTTCA 110  
Qy 21 ThrCyAaGSerAenValIleAaPheAenLeuPheLygInValITyrAaPheAenIleLeu 40  
Db 111 ACTTGTAGAACCAATGTAATGATGACAAATTTATTCMAACAGTTATATATATTTCTT 170  
Qy 41 GluGInGluPheAlaHISaPheGluAlaITyrLeuSerITyrLeuSerITyrLeuSerIleGlu 60  
Db 171 GACCAAGAAATTTGCTCATGATTTTCAAGCTTATCTTTCTTATTGACAAAATATTGAA 230  
Qy 61 SerAenAaenIleAaPheValAaPheLygAaenIleLygValIleAaenValleuSer 80  
Db 231 AGCAACAATAATATTGACAAAGGTTGTAATAATGGAATTAAGTATTAACTTACCTAGC 290  
Qy 81 PheGlyAlaLygGlyAaPheLygSerITyrAaPheAenIleAaPheGluGlnAlaITrPhe 100  
Db 291 TTTGAGGCTTAAGGAGTGAAGAAAACATATGATTAATTCATTGACGACCAACATGGAAT 350  
Qy 101 GluAlaCySerSerAaIThrProValGlnPheValProLyAaenLyAaenITyrLeu 120  
Db 351 GAAGCATGTTCTCTAGAACACCTGTTCAATTTGGTTCTTAAAAACAAGAAATTTATCTT 410  
Qy 121 LeuLygGlnIleIThrPheSerGlyProCyAaGSerSerIleSerValLygIlePheGly 140  
Db 411 CTCAGCAAAATCACCTTTTCAAGTCCATGCAATCTTCTATTTCAATTAAGATTTTGGGA 470  
Qy 141 SerLeuGluAlaSerSerITyrIleSerAaPheITyrAaPheAaITrPheIleAaPhe 160  
Db 471 TCCCTTAGAAGCATCTAGTAAATTTTCAGCTCAAGAAATGAAGGCTTGGATTCCTTTT 530  
Qy 161 AaPheValGlnAaPheValValGlyGlyGlyIThrIleAaenGlyAaenGlyGlnVal 180  
Db 531 CATAGTGTTCAAATTTTATGTTGAGAGAGAGAACTATCAATGACCAATGACAAAGTA 590  
Qy 181 TrpITrProSerSerCyLygIleAaenLygSerLeuProCyAaGAAaPheAlaITrThrala 200  
Db 591 TGGTGCCCAAGTTCTTGGCAAAATTAATAATCATGTCATGACGGAATGCCAACGAGCC 650

Qy 201 LeuThrPheITrAaenCyLygAaenLeuLygValAaenAenLeuLygSerITyrAaenIleGln 220  
Db 651 TTAACCTCTGGAATTTGCAAAATTTGAAGAATTAATCTTAAGAATTAATGACAA 710  
Qy 221 GlnIleHISaIleLygPheGlySerCyAaPheAenValAaSerAaenLeuMetIleAaen 240  
Db 711 CAAATTCATTAATCAATTTTGAAGTCAATGATATGTTAGCTTCAAAATTTGATGATCAT 770  
Qy 241 AlaSerAlaLygSerProAenIThrAaPheLygValHISaValSerAaenIThrITyrIleGln 260  
Db 771 GCTTACGCAAGAGCCCAAAATCTGATGAGTCCAGTATCAATCAATCAATATATTTCAA 830  
Qy 261 IleSerAaPheITrIleIleGlyIThrGlyAaPheCyIleSerIleValSerGlySerGln 280  
Db 831 ATATCTGATTAATTAATTTGAACAGATGATGATTAATTTCAATTTGTTCTGAACTTCA 890  
Qy 281 AaenValGlnAlaIThrAaenIleIThrCyGlyProGlyHISaIleSerIleGlySerLeu 300  
Db 891 AATGTCAGGCGCAAAATTTACTTGTGTCAGGTCATGCTATTAAGTATTGGAAGCTTA 950  
Qy 301 GlySerGlyAaenSerGluAlaITyrValSerAaenValIThrValAaenGluAlaLygIle 320  
Db 951 GGATCTGGAATTTCAAGAGCTTATGCTATATGTTACTGTAATGAAGCCAAATTTATC 1010  
Qy 321 GlyAlaGluAaenGlyValITrGlyIThrGlnITrPheGlnGlySerGlyGlnAlaSerAaen 340  
Db 1011 GGTGCCGAAATGAGGTTAGATTAAGCAAGCTTGCGAGGAGATCTTGACACAACTAGCAAC 1070  
Qy 341 IleLygPheLeuAaenValGluMetGlnAaPheValLygITyrProIleIleIleAaPheAaen 360  
Db 1071 ATCAATTTCTGAATGATGGAATGCAAGCGTTAAGTATCCATATTAATTAACCAAAAC 1130  
Qy 361 TyrCyAaPheAaGlyGluProCyIleGlnGlnPheSerAlaValGlnValLygAaenVal 380  
Db 1131 TATGTGATTCGAGTTGAACCATGTAACCAAGTTTTCAGCGTTCAAGTGAATAATGTG 1190  
Qy 381 ValITyrGluAaenIleLygGlyIThrSerAlaITrITrLygAlaIleLygPheAaPheCySer 400  
Db 1191 GTGATGAAATATCAAGGCGCAAGTGCACAAAGGAGGCGCAATTAATTTGATTTGACAC 1250  
Qy 401 ThrAaPheProCyGluGlyIleIleMetGluAaenIleAaenLeuValGlyGlySerGly 420  
Db 1251 ACAACTTTCATGAGGAAATTAATTAATGAGAAATTAATTTAGAGGAAAGTGA 1310  
Qy 421 LygProSerGluAlaIThrCyLygAaenValHISaPheAaenAlaGlnHISaValIThrPro 440  
Db 1311 AAACCATCAAGAGCTTACGCAAAATGTCATTTTAACAATGCTGAACATGTTACACCA 1370  
Qy 441 HISaIThrSerLeuGluIleSerGluAaPheGluAlaLeuLeuITyrAaenITyr 457  
Db 1371 CACTGCACTTCACTAGAAATTTCAAGAGTGAAGCTCTTTGTATTAATTAAT 1421

RESULT 7  
AR364905 1636 bp DNA linear PART 03-SEP-2003  
LOCUS  
DEFINITION Sequence 2 from patent US 5447867.  
ACCESSION AR364905  
VERSION AR364905.1 GI:34428082  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
REFERENCE  
1 (bases 1 to 1636)  
Bridges, I., Schuch, W. and Grierson, D.  
AUTHORS  
Recombinant DNA containing pectin esterase gene segments  
TITLE  
JOURNAL  
Patent: US 5447867-A 2 05-SEP-1995;  
FEATURES  
source  
Location/Qualifiers  
1..1636  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:



Pred. No.: 6,33e-192 Length: 1636  
 Score: 2390.00 Matches: 457  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-691-374-2 (1-457) x AR364905 (1-1636)

QY 1 MetValIIleGlnArgAsnSerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSer 20  
 Db 51 ARGSTTATCCAAAGAAATAGTATCTCTTCATTATTTTGGTTCATCATTTCA 110  
 QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheLysGlnValTyrAspAsnIleLeu 40  
 Db 111 ACTGTAGACGAAATGTTATTGATGACAAATTTTCAACAAGTTTATGATTAATATCTT 170  
 QY 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerLysAsnIleGlu 60  
 Db 171 GAACAAAGAAATTTGCTCATGATTTTCAAGCTTATCTTCTTATTTTGACCAAAATATTGAA 230  
 QY 61 SerAsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSer 80  
 Db 231 AGCAACATATATATTTGACAAAGTTGATTAAGGATTAAGGATTAATGACTTAC 290  
 QY 81 PheGlyAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGlnGlnAlaTyrPhe 100  
 Db 291 TTTGAGCTTAAGGCTGATGGAACAAATATGATTAATTTGATTCAGCAAGCATGAAAT 350  
 QY 101 GluAlaCysSerSerArgThrProValGlnPheValAlaProLysAsnLysAsnTyrLeu 120  
 Db 351 GAACATGTCATCTTAACAACCTGTTCAATTTTGCTTCTTAAACAAAGATTTATCTT 410  
 QY 121 LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValIlePheGly 140  
 Db 411 CTCAAGCAATTCACCTTTTCAAGTTCATGCAATCTTCTTATTTTCAATTAAGATTTTGA 470  
 QY 141 SerLeuGlnAlaSerSerLysIleSerAspTyrLysAspArgLeuTyrIleAlaPhe 160  
 Db 471 TCTTGAAGCATCTGTAAATTTTCAACTCAAGATGAAGGCTTTGATTCGTTTT 530  
 QY 161 AspSerValGlnAsnLeuValAlaGlyGlyGlyTyrIleAsnGlyAsnGlyGlnVal 180  
 Db 531 GATAGGTTCAAAATTTAGTTGAGAGAGAGAGAACTATCAATGAGCAATGACAAAGTA 590  
 QY 181 TyrTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaProThrAla 200  
 Db 591 TGGTGGCCAAAGTTCTTGCAAAATTAATATCATCGCATGACGAGATGCCAACGCGCC 650  
 QY 201 LeuThrPheTrpAsnCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsnAlaGln 220  
 Db 651 TTACCTTCTGAAATTTGCAAAATTTGAAAGTGAATTAATCTAAGATTAAGTCAACA 710  
 QY 221 GlnIleHisIleLysPheGluSerCysThrAsnValAlaAlaSerAsnLeuMetIleAsn 240  
 Db 711 CAATTCATATCAATTTGAGTCATGCACTAATGTTGAGCTTCAATTTGATGATTCAT 770  
 QY 241 AlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260  
 Db 771 GCTTACGAAAGAGCCAAATATCTGATGAGTATGATCAATTAATCAATATATTCAA 830  
 QY 261 IleserAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280  
 Db 831 ATATCTGATCTATATTGGAACAGTATGATTTTCAATTTGTTCTGAGTCTCAA 890  
 QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300  
 Db 891 AATGCGAGGCGCAAAATATCTGATGAGTATGATCAATTAATTAATTTGAGAGCTTA 950  
 QY 301 GlySerGlyAsnSerGlyAlaTyrValSerAsnValThrValAsnGlnAlaLysIleIle 320  
 Db 951 GGATCTGGAAATTCAGAACTTATGTCTATATGTTACTGTAATGAAGCCAAATTAATC 1010

QY 321 GlyAlaGluAsnGlyValAlaArgIleLysThrTTPGInGlyLysSerGlyGlnAlaSerAsn 340  
 Db 1011 GGTGCGAAATAGAAATGATTTGATGATCAAGACTTGCGAGGAGATCTGCAACAAGTACAC 1070  
 QY 341 IleLysPheLeuAsnValGluMetGlnAspValLysTyrProIleIleIleAspGlnAsn 360  
 Db 1071 ATCAATTTCTGAATGTGAAATGCAAGCGTTAAGTATCCATATTTATGACCAAAAC 1130  
 QY 361 TyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValLysAsnVal 380  
 Db 1131 TATTGATGAGGTTGGAACCATGTATACACAGTTTTCAGCAGTTCAAGTAAAAATGTG 1190  
 QY 381 ValTyrGluAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAspCysSer 400  
 Db 1191 GGTATGAGAAATATCAAGGCGACAAAGTCAACAAGGTGGCCATTAATTTGATTTGACGC 1250  
 QY 401 ThrAsnPheProCysGlyGlyIleIleMetGluAsnIleAsnLeuValGlyLysSerGly 420  
 Db 1251 ACMAACTTTCATGTGAAGAAATTAATGAGATATTAATTTAGTAGGGAAGTGA 1310  
 QY 421 LysProSerGlnAlaThrCysLysAsnValHisPheAsnAsnAlaGlnHisValThrPro 440  
 Db 1311 AAACCATAGAGGCTGACGCAAAAATGTCCATTTTAACATGCTGAACATGTTACACCA 1370  
 QY 441 HisCysThrSerLeuGluIleSerGluAspGluAlaLeuLeuTyrAsnTyr 457  
 Db 1371 CACTGCACTCATAGAAATTTGAGAGATGAAGACTCTTTTGTATATTTAT 1421

RESULT 8  
 LEBOLYGA  
 Locus  
 DEFINITION Tomato gene for cell wall degrading polygalacturonase.  
 VERSION X14074.1 GI:19305  
 KEYWORDS polygalacturonase.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum

REFERENCE  
 1 (bases 1 to 7456)  
 Bird,C.R., Smith,C.J., Ray,J.A., Mourau,P., Bevan,M.W., Bird,A.S.,  
 Hughes,S., Morris,P.C., Grierson,D. and Schuch,W.  
 The tomato polygalacturonase gene and ripening-specific expression

TITLE  
 in transgenic plants  
 JOURNAL  
 Plant Mol. Biol. 11, 651-662 (1988)  
 FEATURES  
 Location/Qualifiers

source

1. 7456  
 /organism="Lycopersicon esculentum"  
 /mol\_type="genomic DNA"  
 /cultivar="Alba Craig"  
 /db\_xref="taxon:4081"  
 /clone\_11b="lambda EMBL3"

repeat\_region

877..904  
 /note="direct repeat 1"

repeat\_region

908..935  
 /note="direct repeat 1"

TATA\_signal

1326..1335  
 /note="direct repeat 1"

CAAT\_signal

1342..1347  
 /note="direct repeat 1"

CAAT\_signal

1400..1403  
 /note="direct repeat 1"

Precursor\_RNA

1411..7244  
 /note="primary transcript"

exon

1411..1757  
 /number=1

TATA\_signal

1431..1435  
 /note="direct repeat 1"

CDS

join(1479..1757,2416..2547,3327..3491,3696..3716,  
 4260..4467,4567..4648,5602..5710,6139..6255,6788..7048)

/product="polygalacturonase"  
 /protein\_id="CAA32235.1"  
 /db\_xref="GI:295813"  
 /db\_xref="GOA:P05117"  
 /db\_xref="UniProt/Swiss-Prot:P05117"



```

/translation="MVIOIRNSILLIIIPASSISICRSNVID)NLKFOVNDILIEOF
AHDROXYLSYKTESNNNI.DKUDKNGIKYIINULSFGAKDGKTYNDIAEQAOMNEA
CSHRTPOQVVPKKNKNTLLKQITRSGPCRSSI SVKIFESLEA;SRISDYKORRLMIAR
DSVQNLVVGGGGTINGNGQVWPSSCKJNKLPCPDAPALTATPMCKNLKYNLKSXN
AQOQHIKRESCTNVASNMIMINASAKSPNTGVAHSNTOYIQISDTIIGTDDCISIV
SGSNVOAKNTICGPHGISIGLSGNSBAVSNVTNEAKIIGAENGVILKTWOGG
VAGSNIKELVNEODVKYPIIIDONCYDRVEPCIQOPSAGVOVKVVENIKGSATK
VAIKPDCSTNPEBGIIMENINLVGSKSPBEATCKNVPNNAEHVTHCTSLSESD
EALVINY"
intron      1758..2415
            /number=1
exon        2416..2547
            /number=2
intron      2548..3326
            /number=2
exon        3327..3491
            /number=3
intron      3492..3695
            /number=3
exon        3696..3716
            /number=4
intron      3717..4259
            /number=4
exon        4260..4467
            /number=5
intron      4468..4566
            /number=5
exon        4567..4648
            /number=6
intron      4649..5601
            /number=6
exon        5602..5710
            /number=7
intron      5711..6138
            /number=7
exon        6139..6255
            /number=8
intron      6256..6787
            /number=8
exon        6788..7244
            /number=9
polyA_site  7244
ORIGIN
polyA_site
Alignment Scores:
Pred. No.:      1,31e-116      Length:      7456
Score:          1506.00        Matches:      454
Percent Similarity: 24.43%    Conservative: 0
Best Local Similarity: 24.43% Mismatches:      3
Query Matchn:   63.01%       Indels:      1404
DB:             8             Gaps:      8
US-10-691-374-2 (1-457) x LEPOLYGA (1-7456)
QY      1 MetValIIegInARpAsnSerIIeLeuLeuIIeIIeIIePheAlaSerSerIIeSer 20
Db      1479 ATGGTTATCCAAAGGAATGATATTCCTCTTCATATATATTTGGCTTCATCAATTTCA 1538
QY      21 ThrCysARSerSerValIIeAspAspAsnLeuPheValIyLeuSerTyLeuSerIyAsnIIeLeu 40
Db      1559 ACTTGTAAAGCAAGTATATGATGACAAATTTATTCAAACAGTTTATGATATATTTCTT 1598
QY      41 GIUGInIuPheAlIHIIeAspPheGlnAlaIyLeuSerTyLeuSerIyAsnIIeLeu 60
Db      1599 GAACAAAGATTTGCTCATGATATTTCAAGCTTATCTTTCTTATTTGAGCAAAATATTTGA 1658
QY      61 SerAsnAsnAsnIIeAspIyValIAspIyAsnGIyIleIyValIIeAsnValIIeSer 80
Db      1659 AGCAACAATTAATATTTGACAAGGTGTATTAATAATGGGATTAAGATTAATGATTAAC 1718
QY      81 PheGIyAlaIyGIyAspGIyIySerThyIyAspAsnIIe----- 93
Db      1719 TTTGAGAGCTAAGGGTGAAGGAAACATATGATTAATATTTGTAAGTATTTAAATATTGGA 1778

```

```

QY      93 ----- 93
Db      1779 TATATTTGTGGGAGTGAATAATGATAGAGATATTAAGATTTATTGGAAAGATGAAGATT 1838
QY      93 ----- 93
Db      1839 ATATTTTATTAAGTGAATAATTTATTTCTCGTTTTAGATTAAGTGAATAATGAGTTTC 1898
QY      93 ----- 93
Db      1899 TCGTTAAGCAGAGAAAGCTATTTTCCATGGTAACGTATTTTTTTTACTTTTAATTA 1958
QY      93 ----- 93
Db      1959 CGTCATGATTTTGTCTATGATCAAGATTAAGACACTATTTATGATGATTAAGTCTCGA 2018
QY      93 ----- 93
Db      2019 AAAGAAATGATGATATTTTGTCTTAATTAACATCAATTTCTTATATGATATTTTTC 2078
QY      93 ----- 93
Db      2079 AACCAAAATTAACAAAGCGTAATCAATTAAGGCGCTAGAAATTAAGATTAAGTTCTAT 2138
QY      93 ----- 93
Db      2139 TCAATTTTAACTTATTTAATTTAGTGAAACCTCGACAAAACGAACAAAGCTATTC 2198
QY      93 ----- 93
Db      2199 AAACCTTTATATTCGGAATTCAGACCAACCATATGAACAACCTGACATGATATAGT 2258
QY      93 ----- 93
Db      2259 CCTAATATATTAATTTTCTTAATAAATACTTCAATCTACCATATTTGAATATTGAAA 2318
QY      93 ----- 93
Db      2319 ATGACTTTATTCCTATCGAACACATATCAAGAGTTTCTTTTAAGAATTTACCACCTACAT 2378
QY      94 ----- -AlaPheGIUGInAlaIyAsnGI 101
Db      2379 TTGATATGTTCTTATCGGTTAATAATATATCTTCAGGATTTGAGCAAGATGAATGA 2438
QY      101 uAlaCysSerSerArgIhrProValGIhPheValIAlProIyAsnIyAsnTyIleuLe 121
Db      2439 AGCATGTTCACTAGAACACCTGTTCAATTTGTGTTCTTAATAAACAAGAAATTAATCTTCT 2498
QY      121 uIyGInIIeThrPheSerGIyProCysArgSerSerIIeSerValIy----- 137
Db      2499 CAAGCAAAATCACTTTTCAGGTCATGACATCTTCTATTTCAAGT-AAAGGTTAGCATAT 2557
QY      137 ----- 137
Db      2558 TGATTTATTAATTCCTCTTTGTAGACAATATATATCTGCTTTATGACAAAATTTAAGA 2617
QY      137 ----- 137
Db      2618 AGTATCAAGATAGATTAACAATGAATTTTGTCACTAATTTAGCGGATTAAGAGAA 2677
QY      137 ----- 137
Db      2678 TTATCAAAATGTATGTTAGTATGAGCAACTTAAGTAAATTAAGTAAAGAAAGTT 2737
QY      137 ----- 137
Db      2738 TGATGCTAATCTATTTTTTTTTTTGTAGAGTAAGATATTTGAACAACATGATTAATTA 2797
QY      137 ----- 137
Db      2798 TTAATTAATGCTTAATTAATATATGATCAATGATAGTCAAACTTAAGAACTGTCAAAAGAA 2857

```

```
QY 137 ----- 137
Db 2858 AATAAGAAAGAAATATTTTATTTTAAATAAATAAAGAAAAATATGAAATTAATT 2917
QY 137 ----- 137
Db 2918 CAAAGCGAGAGATTAATCAATATCTATGCGGATTAAGATATTAATATATGTAAGAAA 2977
QY 137 ----- 137
Db 2978 CAGCACTACACATATCTAATAAGCTCATTAAGATATTAATAAATAAGTGTAGCAAA 3037
QY 137 ----- 137
Db 3038 CAGTTATCCCTACAAAACTTTGTGCGGTAGATCGATCCAGAGTTGTTCCAGACTCT 3097
QY 137 ----- 137
Db 3098 TGCTTAATAAATAATGTTTTTTCTAATAATAGTTGAAAGAAATGTATATGATGAATAAT 3157
QY 137 ----- 137
Db 3158 GAAAGAAAAACATATCAATATTAATAAATAAAGTAATCAAGTAACGAATAACAAATA 3217
QY 137 ----- 137
Db 3218 GGAATAACTCATTAATAAGAAATTTAGTGTTCGTTAACAATACTTAGTTAATTC 3277
QY 138 ----- 141
Db 3278 ATTGTTCTTAAATTCCTTCCTTATTTTGAATAATTAATCAAGATTTTGGATC 3337
QY 141 ----- 161
Db 3338 CTTAAGAGCAATCAAGTAATAATTTTCACTACAAAGATAGAGCTTTGATTCCTTTCA 3397
QY 161 ----- 181
Db 3398 TACTGTTCAAAATTTAGTTGTTGAGAGAGAGAACTATCAATGCGAATGCAAGTATG 3457
QY 181 ----- 192
Db 3458 GTGGCCAAAGTTCTTGCAAAATTAATAATCACTGGTAATTTTAACCTTGCTTAATAGT 3517
QY 192 ----- 192
Db 3518 TTACGCTATGTTGCTCGAATTCCTTAACCTGTTCTAAGATATATATTTGAAGGA 3577
QY 192 ----- 192
Db 3578 GGTGTACAAATGCATCATTTTAAAGATTCCGACCAATATTAGTTTATGTAATCTA 3637
QY 193 ----- 193
Db 3638 ATTTTCAGAGCATCTTTCCTTGAATCATGATCTTACCTTTTTCCTTCATGACGCC 3697
QY 193 ----- 199
Db 3698 ATGCGAGATGCACCAACGTAACGTAATTCATTTGATTAATAAAGGCTTA 3757
QY 199 ----- 199
Db 3758 AAATATATTGAATTTAATTGAAGGTATTAATTAATTTTAACCTTGGCAGAGACTAT 3817
QY 199 ----- 199
Db 3818 TACCCCTTGACATTTAATAGTATTTTAAGATATTAAGTGTTAGTGAACAA 3877
QY 199 ----- 199
Db 3878 AATTAGATATTCAAAACTATTTGAATAATTAATTAATGCAATTTTTCATATCA 3937
QY 199 ----- 199
```

---

```
Db 3938 ATATGATTAATAAATAATTAAGTTAAAGTTCTTAATGATTTGATTCATAAATAAATCAATG 3997
QY 199 ----- 199
Db 3998 ACAAAACAATAGTAGACGGAAGAAATATTAACAATACCTCTTCAAGTAGAATTCATTTGT 4057
QY 199 ----- 199
Db 4058 ACACACACTCAAAAACCTACGTTTCTTGATTAATTAATTCCTTAATTTCTTTAATAGTAA 4117
QY 199 ----- 199
Db 4118 TCAAAAGCTATTAGTCTGTCAAAATCTATACATTGAAACTATCTTGACCCCTGT 4177
QY 199 ----- 199
Db 4178 ACATTGAGATCGTTGAACAAATGATGATTAATTAATTTGATTTAATAATTTAAA 4237
QY 200 ----- 212
Db 4238 ACTAATATTGTTAATTTTCAAGCCTTAACCTTCTGAAATTCGAAATTTGAAAGTGA 4297
QY 212 ----- 232
Db 4298 TAACTTAAAGTAAAAATGACAACAAAATTCATCAATTTGAGTCATGCACTAATGT 4357
QY 232 ----- 252
Db 4358 TGTAGCTTCAAAATTTGATGATCAATGCTTACGAAAGGCCCAAAATGATGAGATCCA 4417
QY 252 ----- 268
Db 4418 TGAATCAAAATCAAAATATATCAAAATCTGATTAATTAATGAACGTTAATTAT 4477
QY 268 ----- 268
Db 4478 TTAATTTTATTTATCCAAATTAATTAAGAAAAAAGAGATTTTATTTGATTAATCA 4537
QY 269 ----- 279
Db 4538 AATTATTAATTTTAAATTTTATTTTATGATGATGATGATTTCAATTTCTTGATC 4597
QY 279 ----- 296
Db 4598 TCAAAATGCGAGGCCCAAAATATTAATCTGTGTCAGGTCAATGATTAAG- GTAATCTTA 4656
QY 296 ----- 296
Db 4657 TTTTACAATATATCTGTTTCATTTTCTATTTTCAATAAAGATATGATATATAA 4716
QY 296 ----- 296
Db 4717 TTACTTTAATCCTTAATTAATTAATTTATTTGCAAAATTTTCTCTTATGATGATAT 4776
QY 296 ----- 296
Db 4777 GACTTAGCAATTAATTAAGGCGGTTTGATGGCGAAATAAAGACGTTTAAAAAGTA 4836
QY 296 ----- 296
Db 4837 CTTTAAAGGTGGAACCTATTTTAAATAAATAGAGTTATCGGTTTGATTAAGTGC 4896
QY 296 ----- 296
Db 4897 TGAAGTTTATGTCAAAACGTGAAGGAAGGAATAAGAAATAAGTTAGGTTATA 4956
QY 296 ----- 296
Db 4957 TGGGTTATTGTATTAATAATTAATTAAGACAAATAAAGATTAATAATGTGTCAACTTAACA 5016
QY 296 ----- 296
```

Db 5017 ACTTAAAGCTACCCCTACCCCTACCCAGCTTTTAACCTTTGGCTTAAATTAAGTTT 5076  
Qy 296 ----- 296  
Db 5077 TTTTAAAACTTAAATAAGTTGTTTGAATATGCGCAAGAGCTAAATATGCAAAAACC 5136  
Qy 296 ----- 296  
Db 5137 AGCTTTAAGTCAGTTTGACCAAGCTTTTAAAGCTGAGCCAAACAGGCTTTAAATATGCTG 5196  
Qy 296 ----- 296  
Db 5197 CTTAAGTGTCTATATATATTTGAGCTTTTGAAGTAGTATATTCCTTAAGTTCAA 5256  
Qy 296 ----- 296  
Db 5257 CATAAATAATCATGCTTTTAACATAGCAATATAGTTAATCAAAAAGCAATATGATGA 5316  
Qy 296 ----- 296  
Db 5317 TTTTGGCAATTTGATTAATTCACAGAAAAAGGATAGTTCAAAAGTACATTCATGAAT 5376  
Qy 296 ----- 296  
Db 5377 TGAAGATATCATAAAGACTAAATTAAGAAATCATATATGAGGATCAAAAATGTTAT 5436  
Qy 296 ----- 296  
Db 5437 TACCTTATTAATAATCTATTCATTTTCATATTAATTAATTAAGAGTGTATTATA 5496  
Qy 296 ----- 296  
Db 5497 ATCTAATAAACAATCAATATTAATTAATGCAAAATGCTTTGGTACCTAATATCTTCT 5556  
Qy 297 ----- 301  
Db 5557 GAATATTTGCTCTATTTTCTCTTTTATTTTTCATGATTTACTATTTGAAAGCTTAGG 5616  
Qy 301 ySerGIyAnSerGIuAlaTyrValSerAnValThrValAangIuAlaIleIleGI 321  
Db 5617 ATCTGAAATTCAGAAAGCTTATGTCTATGTCTACTGTAAATGAAGCAAAATTAATCGG 5676  
Qy 321 yAlaIuAnSerGIyValArgIleIleThrIlePheGln----- 332  
Db 5677 TGCCGAAATGAGTTAGGATCAAGACTTGCA-GGTACCCTCCCCCCCCCCCCCCC 5735  
Qy 332 ----- 332  
Db 5736 CACAGGCCATTTTAAATTTTAAATTTTAAATTTTAAATTCGAATATCAATTAAGATTAA 5795  
Qy 332 ----- 332  
Db 5796 ATTGATTTTCATGTTGAATTTAATTTTGAATTAAGATATGATTTTAACTAGCTTCTAT 5855  
Qy 332 ----- 332  
Db 5856 GTTATATGAAAAAATATGTTCAAGAACTTCAGATTAATGTACTGTAAGTTAAAT 5915  
Qy 332 ----- 332  
Db 5916 GTGTTGCTTTGTTTGAAGTTTGGTTTATCCAGTTTGGGTGATGATTAACCAACTTA 5975  
Qy 332 ----- 332  
Db 5976 TAAATGAAAAAGGGCTGCAACGGCCGACCTAGTCTAGTATCAATAGAAAGATCTCAC 6035  
Qy 332 ----- 332  
Db 6036 GTCTGTTATTCAGATGACGTTCTTGTTGAATGTTAATTAATTAATTAATTAACA 6095  
Qy 333 ----- 338  
Db 6096 TGTATTAAGCATTAATATAATTAATATGCTTTAATATGTAAGGAGATCTGGACAAAC 6155

Qy 338 aSerAnIleIleYsPheLeuAnValGIuMeGIuAnPValIlySerProIleIleIleAs 358  
Db 6156 TAGCAACATCAAAATTTCTGAATGTGAATGCAAGCGTTAAGTATCCATATTAATTAAGA 6215  
Qy 358 pGIuAnSerIlyrCyAsPArGValGIuProCyIleGln----- 370  
Db 6216 CCAAAACTATGTGATCGAGTTGAACCATGTATACACAGGTAAATTTTATTAACGAAC 6275  
Qy 370 ----- 370  
Db 6276 AATTATTATATTTTATTTACTTCTTAATCACCTTACATCATTAACCTTGAGATCTT 6335  
Qy 370 ----- 370  
Db 6336 TTCCTAGTTAGTAACCTTTTGAATAGATTTTAAATGATTAATTCATTAATTCCTTTTA 6395  
Qy 370 ----- 370  
Db 6396 TTTTCTCTAATTTATGATCTTTTGACATATGGCTTAATAATCTGTAAAGTAACT 6455  
Qy 370 ----- 370  
Db 6456 GAATATCATAGAAAAAATGTAGATTAATCTAAATTTTATTAATTAATTAAGAGCTT 6515  
Qy 370 ----- 370  
Db 6516 ATCTAATATTTGTATGTAGATGAGAAACATATATACATTAACATTAATTCATTAATTTA 6575  
Qy 370 ----- 370  
Db 6576 ATTAATATCAAAATTTGATTCGAATCAATCAATCAATCAACAAAGCAATATGAGATT 6635  
Qy 370 ----- 370  
Db 6636 CTCAAAATATAGATGAAGTCAATTCAAAACGAATCAATATATAGATATATCTTAA 6695  
Qy 370 ----- 370  
Db 6696 AAGAGCATTTGGGTAAATTAAGTAAATATCAATTAAGTTAATAAAAAATCTTAACCTGA 6755  
Qy 371 ----- 381  
Db 6756 TCTCTCAGATTAATTAATCACTTGTCTTCAAGTTTTCAGCAAGTTCAAGGAAAAATGTGG 6815  
Qy 381 aIlyrGIuAnIleIlyseGIyThrSerAlaThrIlyValAlaIleIlyPheAspCySerT 401  
Db 6816 TGTATGAAATATCAAGGCAACAGTGCACAAAGGTGGCATTAATAATTTGATTGCAGCA 6875  
Qy 401 hAsnPheProCyAsGIuGIyIleIleMerGIuAnIleAnIleuValGIyGIuSerGIyL 421  
Db 6876 CAATCTTTCATGATTAAGAAATTAATTAATGAATAATTAATTAATTAATTAATTAATTA 6935  
Qy 421 yPProSerGIuAlaThrCyIlyAsAnValIlePheAsnAnIleGIuIleValIThrProH 441  
Db 6936 AACCATTCAGAGCTACGTGCAAAATATGTCATTTTAACATATCTGAACATGTTACACAC 6995  
Qy 441 IeCySerThrSerIleuGIuIleSerGIuAsPArGValIleuIleuIlyrAnIlyr 457  
Db 6996 ACTGCATTCCTCAAGAAATTTCAAGAGATGAAGCTCTTTGTATTAATTAAT 7045

RESULT 9  
TOMPGAAA  
LOCUS TOMPGAAA 7456 bp DNA linear PLN 27-APR-1993  
DEFINITION Tomato polygalacturonase (PG) gene, exons 1-9.  
ACCESSION M37304  
VERSION M37304.1 GI:170472  
KEYWORDS polygalacturonase.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 7456)  
AUTHORS Bird,C.R., Smith,C.J., Ray,J.A., Moreau,P., Bevan,M.W., Bird,A.S.,  
HUGHES,S., Morris,P.C., GRIERSON,D. and Schuch,W.  
TITLE The tomato polygalacturonase gene and ripening-specific expression  
JOURNAL in transgenic plants  
COMMENT Plant Mol. Biol. 11, 651-662 (1988)  
FEATURES  
source Location/Qualifiers  
1..7456  
/organism="Lycopersicon esculentum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4081"  
prim\_transcript 1411..>7456  
/note="polygalacturonase mRNA and introns"  
CDS  
join(1479..1757,2416..2547,3327..3491,3696..3716,  
4260..4467,4567..4648,5602..5710,6139..6255,6788..7048)  
/note="polygalacturonase"  
/codon\_start=1  
/protein\_id="AA34178.1"  
/db\_xref="gi:170473"  
/translation="MVYQRNSILLLIIIFASSISTCRNVIDNLEKQVYDNLIEQEF  
AHDROAVLSTSKTESNNNDIKVDKGIKIVINVLFGAGDKETDNLAEQANMEA  
CSRTPVQFVPEKKNKYLKQITFSGCRSISVIGISLEASKISDYKDRLLMTAF  
DSVNLIVGGGGTINNGOVWPSCKINSLPCRDAPLPTFMKNKNLKNLKN  
AQQIHIRESCITNVASNLIMINASAKPNTDGVHNTVOYIOISDTIGGDCISIV  
SGSQNVATNITCGPHGISIGSLGSGNSFAVSVNTVNAKLTIGENGRIRKTWGG  
SGQSNIKPLNEMQDVKPIIIDQNCDEVEPCIOQFSAVOGVNVAENIKQTSATK  
VAIKFDSTNPCEGIIIMENINLVGSGKFSKATCKNVHNAHTPCHTSLEISD  
EALLNNY"  
exon <1479..1757  
/note="polygalacturonase"  
/number=1  
intron 1758..2415  
/note="polygalacturonase intron A"  
exon 2416..2547  
/number=2  
intron 2548..3326  
/note="polygalacturonase intron B"  
exon 3327..3491  
3492..3695  
/note="polygalacturonase intron C"  
intron 3696..3716  
3717..4259  
/note="polygalacturonase intron D"  
exon 4260..4467  
4468..4566  
/note="polygalacturonase intron E"  
intron 4567..4648  
/number=6  
exon 4649..5601  
/note="polygalacturonase intron F (no splice consensus);  
putative"  
intron 5602..5710  
/number=7  
exon 5711..6138  
/note="polygalacturonase intron G"  
intron 6139..6255  
6256..6787  
/note="polygalacturonase intron H"  
exon 6788..>7048  
/note="polygalacturonase"  
/number=9  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,31e-116 Length: 7456  
Score: 1506.00 Matches: 454  
Percent Similarity: 24.43% Conservative: 0  
Best Local Similarity: 24.43% Mismatches: 3  
Query Match: 63.01% Indels: 1404  
DB: 8 Gaps: 8

US-10-691-374-2 (1-457) x TOMPGAA (1-7456)  
QY 1 MetValIleGlnArgAsnSerIleLeuLeuIleIlePheAlaSerIleSer 20  
Db 1479 ATGGTATCCAAAGAAAGTATTCCTTCCTCATTAATTAATTTGCTCAATCAATTCA 1538  
QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheValIleTyrAspAsnIleLeu 40  
Db 1539 ACTGTGAGAGCAATGTTATGATGACATTTATCAACAAAGTTATGATATATTCCTT 1598  
QY 41 GluGlnGluPheAlaHisAspPheGlnIleTyrLeuSerTyrLeuSerIleGlu 60  
Db 1599 GAACAGAAATTTGCTCATATTTTCAAGCTTATCTTCTTATTTGAAGCAAAATATTTGAA 1658  
QY 61 SerAsnAsnAsnIleAspIleValIleAspAsnGlyIleIleValIleAsnValLeuSer 80  
Db 1659 ACCAACAAATTAATTAATGACAGGTTGATTAATAAGGATTAATGATTAATGCTTAGC 1718  
QY 81 PheGlyAlaIleGlyAspGlyIleThrTyrAspAsnIle----- 93  
Db 1719 TTGGAGCTTAAGGATGATGCAAAACATATATATATTTGAAGTATTAATATTTGAA 1778  
QY 93 ----- 93  
Db 1779 TATATTTGTGGGATGAATAATGATAGAAATATAAGAAATTTTGAAGATGAAAGTT 1838  
QY 93 ----- 93  
Db 1839 ATATTTTAAGATGAATAATTAATTTTCTGTTTATAGATTAAGTGAATAAGATTTTC 1898  
QY 93 ----- 93  
Db 1899 TCGTTAAGCGAAGAAAGTATTTTCCATGTAAGTATTTTCTTACTTTTAATTA 1958  
QY 93 ----- 93  
Db 1959 CGTATAGTATTTGCTACTCAAGATTAAGACATTAATTAATGATTAAGATTTAGCTCGA 2018  
QY 93 ----- 93  
Db 2019 AAGAAATGATATGTAATTTTGTCTTAATATACTCAATTTCTTAATATGATATTTTC 2078  
QY 93 ----- 93  
Db 2079 AACCAAAATACAAAGCGTAATCAATAAGTGGCTCTAGATTAAGATTAAGTCTAT 2138  
QY 93 ----- 93  
Db 2139 TCAATCTTAACCTTAATTTAATTTTAGTGAACCTCCAGCAAAACGAAACGATATTC 2198  
QY 93 ----- 93  
Db 2199 AAACCTTTAATTCGAATTCGAGCAACATATGAAACAACCTGACATGCAATATAGT 2258  
QY 93 ----- 93  
Db 2259 CCTAATATATATATTTTCTTAATAAATATCTTAATCAATATGAAATATGAAAA 2318  
QY 93 ----- 93  
Db 2319 ATGACTTTATTCCTATCGACACATATCAAGAGTTCTTTAAGAAATTTACCACTATCAT 2378  
QY 94 -----AlaPheGluGlnAlaIleTyrPheGlu 101  
Db 2379 TTGATATGTTTCTTATCGTGTGTTAAATTAATTTTCAAGCAATTTGACGAAGATGAAATGA 2438  
QY 101 uAlaCysSerSerArgThrProValGlnPheValIleProIleAsnIleAsnTyrLeuLeu 121  
Db 2439 AGCATGTCATCTGAGAACACCTGTTCAATTTGTGTTCTTAATAAACAAGATTAATCTTCT 2498  
QY 121 uIleGlnIleThrPheSerGlyProCysArgSerSerIleSerValIle----- 137  
Db 2499 CAAAGAAATCACCTTTTCAAGGTCATGCAATCTTCTATTTTCAGT-AAAGTTAGCATAT 2557

QY 137 ----- 137  
Db 2558 TGATATTATTAATCCTTTGTAGCAATATATATCTGGTTATGACAAATTAAGAA 2617  
QY 137 ----- 137  
Db 2618 AGTAATCAAGAATAGATAAACAATGATTTTGTCTACTATATTAGCGGATTAGTGAGAA 2677  
QY 137 ----- 137  
Db 2678 TTATCAAAATGTATAGTAGTATGACAACTTAGCTATGAATAGCTATGAGAAGTT 2737  
QY 137 ----- 137  
Db 2738 TGATCTAATCTAATTTTTTTTTTGTAGAGTAAGATATTGAAAACATGTATTAATTA 2797  
QY 137 ----- 137  
Db 2798 TTAAATTAATCTTAATTAATATGTCAATGATAGTTCAAACTAAGAACTGTCAAAAGAA 2857  
QY 137 ----- 137  
Db 2858 AATTAAGAAAGAAATATTTATTTTAAATAAATTAAGAAAGAAATATGAGAAATTAAT 2917  
QY 137 ----- 137  
Db 2918 CAAGCAGAAAGGTATTACATACTATGCGGATTAAGATATTAATATGTAAAGAAA 2977  
QY 137 ----- 137  
Db 2978 CAGCACTACACATCTAATAAAGTCTCATAAATGATATAAATAATAGTGTAGAACAA 3037  
QY 137 ----- 137  
Db 3038 CAGTTATCCCTACAAAACCTTTTGTGGGTAGATGATCAGAGGTTGTTCAAGCTCT 3097  
QY 137 ----- 137  
Db 3098 TGCTTAAGAAAGTGTTTTCTAATAATGATTGAAGAAATGTATATGATGAAGAAATAT 3157  
QY 137 ----- 137  
Db 3158 GAAAGAAAACATATCAATTTTAAAAATTAATAAGTAACTAAGTAAACAAATTAACAATA 3217  
QY 137 ----- 137  
Db 3218 GGAATATATCTCATTAATGAATAATTAAGGCTTTTCGTTAATATCTTAGTTTATTC 3277  
QY 138 ----- 141  
Db 3278 ATTTGTTCTTAATTTCCCTCTTATTTTTTTTGAATTAATCAATGACAGATTTTGGATC 3337  
QY 141 ----- 161  
Db 3338 CTTAAGAGCATCTTAAGAAATTTTCAGACTTACAAAGATAGAGGCTTTGATTTGCTTTTGA 3397  
QY 161 ----- 181  
Db 3398 TAGTGTCAAAATTTTAGTTGTTGAGAGAGAGAACTATCAATGCAATGCAAGTATAG 3457  
QY 181 ----- 192  
Db 3458 GTGGCCAGATCTTCCAAATTAATTAATCACTGTGTAATTTTATTAACCTTGCTTAATAGT 3517  
QY 192 ----- 192  
Db 3518 TTTAGCTATGTTGCTCGAATCTTTAAACCTGTTCTAAGATATTAATATATTGGAAGA 3577  
QY 192 ----- 192  
Db 3578 GGTTGTCAAAATGATCATCATTTTATAGAGATTCGACCAATATTAAGTTTATGTAATCTA 3637

QY 193 -----Pr 193  
Db 3638 ATTTCAAGAGCATCTTGCCCTTGTAAGTATGATCATTTACCCCTTTTCTTCATGACAGCC 3697  
QY 193 ----- 199  
Db 3698 ATGCAGGATGACCAACGGTACGTTAATGTGATTTGATTAATAAAAAAGCCCTA 3757  
QY 199 ----- 199  
Db 3758 AAATATATTGAATTTTAATGAAAAGTATATATCTTAACCTTTGGGAGAGACTAT 3817  
QY 199 ----- 199  
Db 3818 TACCCCTTGACTAATTAATAGTATTTTAAAGATATAAGTGTAGTGAACAAA 3877  
QY 199 ----- 199  
Db 3878 AATTAGATTTCAAAAACATAATTGAAAATTACTATAAATTGCAATTTTTTGCATATCA 3937  
QY 199 ----- 199  
Db 3938 ATATGATTAATAAATATTAAGTTAAGTCTTATGATTTGATCTTAATAATAAATATCATAG 3997  
QY 199 ----- 199  
Db 3998 ACAAAACAATAGTAGCGAGAGAAAGTATATAACAATACCTTTCAAGTAAATCATTTGT 4057  
QY 199 ----- 199  
Db 4058 ACACACACTCAAAAACCTACGTTTCTTGATTTATATTTCTATTTCTTAATAGTAA 4117  
QY 199 ----- 199  
Db 4118 TCAAGGCTATTAAGTCTGTCAAAATCTATACATTGAAACCTATCTTTGACGCTCGT 4177  
QY 199 ----- 199  
Db 4178 ACATTGAGATCGTTGAACAATGATGAAATGATTAATTAACCTTGTAATTAATTAATA 4237  
QY 200 ----- 212  
Db 4238 ACTAATATGTTTAATTTTCAAGGCTTAACCTTCGAATTCGAAAATTTGAAGTGA 4297  
QY 212 ----- 232  
Db 4298 TAACTTAAGATTAATAATGACACAAATTCATATCAAAATTTGAGTCATGCATTAATGT 4357  
QY 232 ----- 252  
Db 4358 TGTACCTTCAAAATTTGATGATCAATGCTTCAGCAAGAGCCCAATTAATGATGAGTCCA 4417  
QY 252 ----- 268  
Db 4418 TGTATCAATCTCAATATATTTCAATATCTGATATCTATTAATGGAACGGTTATTAAT 4477  
QY 268 ----- 268  
Db 4478 TTAATTTTATTTATCAATTAATTAAGAAAAAGAGATATTTTATTTGATTAATCA 4537  
QY 269 ----- 279  
Db 4538 AATTATTAATTTTAATTTTATTAAGGATATGATTAATTTGATTTGCTGAGATC 4597  
QY 279 ----- 296  
Db 4598 TCAAAATGTGACGACCAATATTAATTAATTTGAGTCAAGTATGATTAAG-CTACTCTA 4656  
QY 296 ----- 296  
Db 4657 TTTTACAATATTAATCTGTTTCATTTCTCTAATTTCAATAAAGTATGATATTAATAA 4716  
QY 296 ----- 296

```
Db 4717 TTACTTAAATCCTTAAATTAATTATTTGCAAAATTTTCTGTGCTTATGCTAAT 4776
QY 296 ----- 296
Db 4777 GACTAGACAAATTAATAGGCGCTTGGATGGCGAATAAAGAGCTTTAAAGTA 4836
QY 296 ----- 296
Db 4837 CTTTAAAGTGTGAAACTTATTTTAAATAGACGTTATCGGTTGATAAAGTGC 4896
QY 296 ----- 296
Db 4897 TGAAGTTGTATGCAAAAGTGAAGGAAATATGAAAGAAATGTTAGGTTATA 4956
QY 296 ----- 296
Db 4957 TGGGTTATTTGTATATAAAATATTAAACAACAAAGATAAATGTGTCACTTAACA 5016
QY 296 ----- 296
Db 5017 ACTTAAAGCTACCTACCTACCCAGCTTTAACTTTGGCTTAAATAGTTTTTT 5076
QY 296 ----- 296
Db 5077 TTTTAAACTTAAATTAAGTGTGTTGAGTATGCCAAGAGCTAATATGCAAAACC 5136
QY 296 ----- 296
Db 5137 AGCTTTAAGTCAGTTTGACACAGCTTTTAAGCTGACCAACAGGCTTTAAATGTCTG 5196
QY 296 ----- 296
Db 5197 CTTAGATGCTATATATATTTAGCTTTTGAAGTAGTATATTCTTAAAGTTCA 5256
QY 296 ----- 296
Db 5257 CATTAATATACGTGTTTAACATAGCACATATAGTTAATCAAAAGACAAATGATGATA 5316
QY 296 ----- 296
Db 5317 TTTTGCGAATTTGATTAATTCACAGAAAGGATAGTTCAAGGTACATTCAATGAT 5376
QY 296 ----- 296
Db 5377 TGAAGATATCATAAAGATTAATTAAGAAATCAATATGAGGATCAAAATGTAT 5436
QY 296 ----- 296
Db 5437 TACTTATTAATACTATTCATTTTCATATTAATTAAGTAAGTGTATTA 5496
QY 296 ----- 296
Db 5497 ATCTAATAAAACATGCAATTAATTAATGACAAATGTGTTGTACTATATCTTCT 5556
QY 297 ----- 116glYserLeuGI 301
Db 5557 GAATATTTGCTATTTTCTCTTTTATTTTCCATGATTAATTAATGAGCTTAGG 5616
QY 301 ----- 116glYserLeuGI 301
Db 5617 ATCTGAAATTCAGAGCTTATGTCTTAATGTTACTGTAATGAGCCAAATTTATCGG 5676
QY 321 ----- 332
Db 5677 TGCCGAAATATGAGTTAGATCAAGACTTGGCA-GTAACTCCCTCCCCCCCCCCCC 5735
QY 332 ----- 332
Db 5736 CACAGGCCATTTTAAATTTTAAATTTTATTCGAATATCAATTTAAAGTTA 5795
QY 332 ----- 332
Db 5796 ATTGATTCATGCTTGAATTTATATTGATTAAGATATGATTTTACTAGCTTCTAT 5855
QY 332 ----- 332
Db 5856 GTTATATAGAAAAAATATGTTCAAGACTTCAGATTAATTGACTGACTAAGTAAAT 5915
QY 332 ----- 332
Db 5916 GTGTGCTTTGTTAGAACTTTGTTTATCCAGTTTGGTGATGATTAACCAACTTA 5975
QY 332 ----- 332
Db 5976 TAATGAAAGGGGCTGCAACGCCGCCCACTAGTCTAGTATCAATAGAAATCTCAC 6035
QY 332 ----- 332
Db 6036 GTCTGTTATTCAGATGACGCTTGTGGTGAATGTAATTAATTAATTAACA 6095
QY 333 ----- 61glYserGIyGIhAI 338
Db 6096 TGTATTAAGCATTAATTAATTAATGTTATGTTTAAATAGTGAAGAGATCTGACAGC 6155
QY 338 ----- 61glYserGIyGIhAI 338
Db 6156 TATGAAATTAATTAATTAATTAATGTTATGTTTAAATAGTGAAGAGATCTGACAGC 6215
QY 338 ----- 61glYserGIyGIhAI 338
Db 6216 CCAAACTATGATGATGAGTTGACCACTATTAACAGGTAAATTTTATTAACGAA 6275
QY 370 ----- 370
Db 6276 AATTATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6335
QY 370 ----- 370
Db 6336 TTCACTAGTTAGTAATTTTGAATAGATTTTGAATTAATTAATTAATTAATTAATTA 6395
QY 370 ----- 370
Db 6396 TTTTCTTAATTTATGATCTTTTGACATAGTCTTAATTAATTAATTAATTAATTA 6455
QY 370 ----- 370
Db 6456 GAATATCATTAAGAAAAATGTTAGATTAATTAATTAATTAATTAATTAATTAATTA 6515
QY 370 ----- 370
Db 6516 ATCTAATTTTGTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 6575
QY 370 ----- 370
Db 6576 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6635
QY 370 ----- 370
Db 6636 CTCAAAATATGATGATGATGATTTCAAAAGAAATCAATATAGTATATCTTAA 6695
QY 370 ----- 370
Db 6696 AAGAGAGATTTGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6755
QY 371 ----- 61glYserGIyGIhAI 381
Db 6756 TCTCTACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6815
QY 381 ----- 61glYserGIyGIhAI 381
Db 6816 TGTATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6875
QY 401 ----- 401
Db 6876 CAACTTTCAATGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6935
QY 401 ----- 401
```



Oy		361	TyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGluValLysAsnVal	380
Dd		1120	TACGTATGCAAAGCAAAACCCTTGTCCAAAGAACAGAGCTCGGCTGTCCAAAGTGAATAATTATT	1177
Oy		381	ValTyrgluAsnIleLysGlyThrSerAlaThrLyValAlaIleLysPheAspCysSer	400
Dd		1180	TTTCACAAGAACATCAAAGCACATGTGCTCTCAAAATGCCCAATTAATTGAATGGACAC	1233
Oy		401	ThrsAsnPheProCysGluGlyIleIleMetGluAsnIleAsnLeuValGlyGluSergly	420
Dd		1240	AAGAGTGTCCTCATGCCAGGAGATTGTGTGGAAGATGTTGATCTAAGATATAAGAGACT	1299
Oy		421	LysPserGluAlaThrCysLysAsnValHisPheAsnAsnAlaGluHisValThrPro	440
Dd		1300	GCAGCAGCAAAAGCTTTTGCAACAAATGTCAACTCAGCAAAAACGAGAGTGTTCACCA	1359
Oy		441	HisCysThrSerLeuGluLysSerGluAspGluAla	452
Dd		1360	CATTGCCCAAGAA-----GGAGAAAGAAAGACT	1386
RESULT 11				
LOCUS	AY043233	1595 bp	mRNA	linear PLN 20-MAY-2003
DEFINITION	Vitis vinifera polygalacturonase mRNA, complete cds.			
ACCESSION	AY043233			
VERSION	AY043233.1	GI:15081599		
KEYWORDS				
SOURCE	Vitis vinifera			
ORGANISM	Vitis vinifera			
REFERENCE	Bukacynska; Vitidiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
AUTHORS	rosids; Vitaceae; Vitis.			
TITLE	1 (bases 1 to 1595)			
JOURNAL	Nunan,K.J., Davies,C., Robinson,S.P. and Fincher,G.B.			
MEDLINE	Expression patterns of cell wall-modifying enzymes during grape berry development			
PUBMED	Planta 214 (2), 257-264 (2001)			
REFERENCE	2 (bases 1 to 1595)			
AUTHORS	Fincher,G.B.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-JUL-2001) Dept. Plant Science, University of Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia			
REFERENCE	3 (bases 1 to 1595)			
AUTHORS	Nunan,K.J., Davies,C. and Robinson,S.P.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-JUL-2001) Plant Industry, CSIRO, Hartley Grove, Adelaide, SA 5064, Australia			
FEATURES	Location/Qualifiers			
source	1..1595			
	/organism="Vitis vinifera"			
	/mol_type="mRNA"			
	/cultiyar="Shiraz"			
	/db_xref="taxon:29760"			
	55..1389			
	/codon_start=1			
	/product="Polygalacturonase pgl"			
	/protein_id="AK81876.1"			
	/db_xref="gi:15081600"			
	/translation="MSIQRRFVCILVFLLSPFCFSLSQASPPDYLGASGYDSQAN			
	HHRSTIQSRFTWMPFRKFSKTGSKLASVKMWNVNYGAGDSDATEAFKAMKA			
	ACSSPGSLVLPNNKYLLKPITFGPCPKSSLTVOIYGVQASTRSAVNDMTFWLI			
	FENVNLAIVOGGTINGNGKTWMNSCKVNYDLPCGAPTLALFPNCKNLAVKNLKIO			
	NADQMHSFEKVQVSGGLTPAPGNSPNTDGIIHVSYTKIIOISSVYIGDCDISI			
	VSSSQKOVNQGITCGPHGISISLGSGPSSEAHVDYTNKGATTEGSTINGVRIRIKTWG			
	GSSASNIKRQNIWMHNENPIITIIDKTYDDGSKPPCKSSSAIVOVNVIYONIKGTSSS			
	KKAISLDCAKKPFCCGILLRIDIDIKVGGSRAKAAVAACSARVYMWDVSNPCA"			
ORIGIN				
Alignment Scores:				
Prod. No.:	5.38e-92	Length:		1595

Score:	1204.50	Matches:	228
Percent Similarity:	75.74%	Conservative:	53
Best Local Similarity:	61.46%	Mismatch:	87
Query Match:	50.40%	Indels:	3
DB:	8	Gaps:	3
US-10-691-374-2 (1-457) x AY043233 (1-1595)			
QY 73	IIleYsValIIleAenValIleuSePheGlyValAlYsgIlybPthThYrAspAsn	92	
Db 277	GTIAAAATGGTTAAACGGAATTATTAATGAGCTTAAAGATGATGATGATGCCACAGAG	333	
QY 93	IIleAlaPheGluGlnAlaTrPAsnGlnAlaCYseSerAerAghThrProValGlnPheVal	112	
Db 337	---GCATTCAAGAAAGCTTGGAAAGCAGCTTGTTCATCTCAGAGATCTGTC---CTAGTG	399	
QY 113	ValProIlyAsnIlybAnTYrLeuLeuIlyeGlnIleThrPheSerGlyProCYaArgSer	132	
Db 391	GTGCGCAAAACAGAACTATCTCTTAAAGCAATTAATCAATTCAGAGGCTCTGTAAATCC	450	
QY 133	SeTIIseRValIlysIIlePheGlySerLeuGlnAlaSerSerIlysIIseRAspTYr---	151	
Db 451	AGATTACAGTGCAGATATATGAGACGTCTCAGACATCTACTATGATGATCGCTACAGT	510	
QY 152	IlyAspArgArgLeuTrpIIleAlaPheAspSerValGlnAsnLeuValAlGlyGlyGly	171	
Db 511	AAAGATATGACCCACACGCGCTTATCTTTGAAAATGTCAAAATTTAGCGGTCAAGGTGT	570	
QY 172	GlyThrIIleAsnGlyAsnGlyGlnValTrpTrpProSerSerCYsIlysIIleAnIlySer	191	
Db 571	GGACACATCAATGGAATGGAAGAGCGTGTGGAAACCTCTTGCAAAAGTGAACATATGAT	630	
QY 192	LeuProCYaArgAspAlaProThrAlaLeuThrPheTrpAsnCYsIlyAsnLeuIlyVal	211	
Db 631	CTTCTCTTGAAGAGGTGACCAACGCGCTTAACCTTCTAATCACTGCAAGACCTTGGCAGTG	690	
QY 212	AsnAsnIleuIlySerIlyAsnAlaGlnGlnIleHisIleIlyPheGluSerCYsThrAsn	231	
Db 691	AAAGAACCTTAAGATCCAAATATGCCAGCAAAAGCATGTTCTTTGAAAAATGCGGTGT	750	
QY 232	ValValAlaSerAsnLeuMetIIleAsnAlaSerAlaIlySerProAsnThrAspGlyVal	251	
Db 751	GTTCAGCTTCGGGTCTCACCATTACTGACCGGGGACAGTCCCAACATATGAAATC	810	
QY 252	HisValSerAsnThrGlnTrpIleGlnIleSerAspThrIleIleGlyThrGlyAspAsp	271	
Db 811	CATGGAGGTACACCAAAATTAATTCAGATTTCAACGTCTGATTATGGAAACAGTATGAC	870	
QY 272	CysIIseSerIIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCYseGlyPro	291	
Db 871	TGTATTTCATTTGTGAGTGGATCCCAAGGTGCAAGTCAATGCTATCACTGTGGACCA	930	
QY 292	GlyHisGlyIIleSerIIleGlySerLeuGlySerGlyAsnSerGlnAlaTrpValSerAsn	311	
Db 931	GGCCATGAAACAGACATTTGGAAGCTTGAAGTCTGGAATTCAGAGGCTCATGTTTCGAT	990	
QY 312	ValThrValAsnGlnAlaIlyeIIleIleGlyAlaGluLeuGlyValArgIIleYsThrTrp	331	
Db 991	GTTCACGTGAACGGCGCAACGCTTTCGGGAACCAAAATGAGATTTAGATTAACAACATGG	1050	
QY 332	GlnGlyGlySerGlyGlnAlaSerAsnIleIlyPheLeuAsnValGluMetGlnAspVal	351	
Db 1051	CAGGAGAGGTCTGGAAGCGCAAGCAACATCAAAATTTGAGAACTGTAAATGCAATACGTG	1110	
QY 352	IlyeYrProIleIleIleAspGlnAsnTrpCYaAspArgValGluProCYsIIleGlnGln	371	
Db 1111	GAGAACCTTAATTAATTAATGACCAAAAGTACTGTGACCAAAAGCAAAACATGACATGTCAG	1170	
QY 372	PheSerAlaValGlnValIlyAsnValValTrpGluAsnIleIlyGlyTrpSerAlaThr	391	
Db 1171	AGCTCGGCTGTTCAGTGAAGAAATGTCTTACCAAGACATTAAGGAACAAGTTCTTCG	1230	
QY 392	IlyeValAlaIleIlyPheAspCYsSerThrAsnPheProCYsGluGlyIleIleMetGlu	411	



```

Db      1231 AAGGAGCCTATATCGCTTATTCAGATCGGAACTTTCATGTCACAGGAGTTCTGCTGCCA 1290
Qy      412 AenlleaenleuValglYgluSerGlYlyProSerGluAlaThrCyAlaYsaenValHis 431
Db      1291 GACATTGATATTAAGTTGAGAGGAAAGGCAAGGCTGTATGACAGCATGCCACA 1350
Qy      432 PheAenAenAlaGluHisValThrProHisCyG 442
Db      1351 GTGACTGTATGAGGGAGGTGTCTCCAAATTCG 1383

RESULT 12
AVOPOLYCAL
LOCUS      AVOPOLYCAL               1725 bp      mRNA      linear      PLN 09-AUG-1994
DEFINITION Avocado polygalacturonase mRNA, complete cds.
ACCESSION  L06094
VERSION     L06094.1 GI:166950
KEYWORDS   pectin depolymerase; pectinase; polygalacturonase.
SOURCE      Persea americana (avocado)
ORGANISM    Persea americana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; magnoliids; Laurales; Lauraceae;
            Persa.
REFERENCE   1 (bases 1 to 1725)
            Kuriunai S.Y., Lin A.C., Percival, F.W., Laties, G.G. and
            Christoffersen, R.E.
            Ripening-related polygalacturonase cDNA from avocado
            JOURNAL Plant Physiol. 103 (1), 289-290 (1993)
            MEDLINE 94269193
            PUBMED  8208850
COMMENT     Original source text: Persea americana (strain Hase) ripe fruit
            mesocarp
FEATURES
    source
        1..1725
            Location/Qualifiers
                /organism="Persea americana"
                /mol_type="mRNA"
                /strain="Hase"
                /db_xref="taxon:3435"
                /tissue_type="mesocarp"
                /dev_stage="ripe fruit"
                1..33
                34..1395
                /FC_number="3.2.1.15"
                /function="cell wall degradation"
                /codon_start=1
                /product="polygalacturonase"
                /protein_id="AA032914.1"
                /db_xref="GI:166951"
                /translation="MALTRLPLPISLWRCFYSSHTLLOKDPILICVNDPRPDQRAY
                PTYFGILDBESSIMGFEPSSLSLEKFNVGPEPDSVDIVDFGARGDGTDTKA
                FEKAWMDACSSSVLIVPENKYLKQITFGSPCKSDLVKIRGTIEASDSDWGH
                NKRRIEPEIDISNLIEGGGTINGNETWDSCKRKKSLPKSAPTALTFPSCKVLI
                VSDLSIKDSOKMHLSPDKCQDIYASNLMTAPESHSPNTGIIHTGRIHWNMSVIGT
                GDDCSIETSGSKMVIATNITCGPHGISLGSIDRNSBAHVSAGVLVDGGLPDTNGL
                RKTQNGSGSKAKITKQNIWANTNPRIITDQYCDSDPPEPDSATKYNVAYMN
                INGTSHSEVAIVKFDCKSPCCGILLENINLVNGGKETTMSCSNIVQGTGKQYPP
                SCL"
                1394..1725
    CDS
        5'UTR
        3'UTR
    ORIGIN
        3'UTR
        1394..1725
    Alignment Scores:
        Pred. No.:      1,3e-87      Length:      1725
        Score:          1153.00      Matches:      225
        Percent Similarity: 72.35%      Conservative: 68
        Best Local Similarity: 55.56%      Mismatches: 104
        Query Match:      48.24%      Indels:      8
        DB:                Gaps:      5
    US-10-691-374-2 (1-457) x AVOPOLYCAL (1-1725)
    Qy      28 AspAspAenleuPheYsaenValYrAspAenlleuGlGluGlnIuPheAlaHisAap 47
    Db      151 GACCAAGAGGCCCTACCCCACTATTTTGGTCCATTAGAT---GAGTTCTCTACATC 207

```

```

Qy      48 PheGlnAlaTYrLeuSerTYrLeuSerlysaenlleGluSerAasnAasnleAaplyS 67
Db      208 ATGGAGTTTGAGCCCTCCATCTTGAGT-----TTGGAGCGGTTTAACTCAATGGGTGT 261
Qy      68 ValAaplysaenGlyYlleYsaVallleaenValleuSerPheGlyValaYsaGlyAapGly 87
Db      262 CCGGAACCTTCCACAGATACAGATATTAAGTGTGATGATTTTGGAGCCAGAGGTGATGG 321
Qy      88 lyerThYrAspAenlleaPheGluGlnAlaTPaenGluAlaCySerSerAargThr 107
Db      322 ---ACAGATGACACAAAGGCATTGTGAGAGCGCTTGGAGAGCGCTGTTCATCAGATCT 378
Qy      108 ProValGlnPheValYalProlysaenlysaenTYrLeuLeuYsaGlnleThrPheSer 127
Db      379 GTCCCTC-----ATTGTGCTGTAAACAGAACTATCTTAAACAAATTAACCTTTTCA 432
Qy      128 GlyProCyArgSerSerlleSerVallySillePheGlySerleuGlnAlaSerSerlyS 147
Db      433 GGTCTTGTAAATCTGATCTTGTGTGAGATTCGCGGACCATTTGACCTTCATCTGAT 492
Qy      148 lIeSerAspTYr-----lysaPargArgleuTPilleaPheAapSerValGlnaen 165
Db      493 CATCAGATTTGGGTGGACATATATCGAATAACGATGATTTGAATTTGAACATCAGCAAC 552
Qy      166 leuValValGlyGlyGlyYThrIleAenGlyAenGlyValYrTPProSerSer 185
Db      553 CTCACATGAAAGGGGGGTGAACCATCAATGGAATGAGAGAACATGGGGATAGCTCC 612
Qy      186 CylySilleAasnlySerleuProCyArgAapAlaProThAlaLeuThrPheTPaen 205
Db      613 TCCAAAAGGAAGAAATCCCTTCTTGCAAAAGCGACCTACGGCTTAACCTTCAGCTGC 672
Qy      206 CylysaenleuValaAasnAenleuYSerlysaenAlaGlnIleHisIlelyS 225
Db      672 TCCAAAGAACCTGATTTGTGAGCATCTCAGTATCAAGACAGCAAGCAAAAGATGATCTTCT 732
Qy      226 PheGluSerCyThraenValYalAaSerAenleuMeIleAasnAaSerAlaYsaer 245
Db      733 TTTGATTAAGTGCAGATGTCAATGCTTCAATGATGATGATGCAAGACATAC 792
Qy      246 ProAenThraSPGlyValHisValSerAenThraGlnTYrIleGlnIleSerAapThrIle 265
Db      793 CCAACACAGATGGCATTTATATTAACAGGACCCAAAGAAATCAATGATGATTCGTC 852
Qy      266 lIeGlyThraYsaSPCySilleSerIleValaSerGlySerGlnAasnValGlnAlaThr 285
Db      853 ATCGGAACAGGTGATGATGTATCTCAATAGAGATGATCAAAAATGATGATAGCCACA 912
Qy      286 AenlleThrCyGlyProGlyYHleGlyYlleSerIleGlySerleuGlyAenSer 305
Db      913 AACATAAACCTTGAGACCAAGCCATGAAATTAACATTGAAAGCTTGAAGATCGAAATTTCA 972
Qy      306 GluAlaTYrValSerAasnValThrValaenGluAlaYlelleGlyAlaGluAenGly 325
Db      973 GAAGCTCATGTTTCAGAGTACTTGTGATGGCGGCAACCTTTGCATACAAATATGA 1032
Qy      326 ValArgIlelySerThrTPGlnIlyGlySerGlyGlnAlaSerAasnIleYsaPheleuAen 345
Db      1033 CTCAGATCAAAAACATGGCAGGAGAGGTTCAAGAGTCAAGAAATCAATCAAAATC 1092
Qy      346 ValGluMetGlnAapValYrTYrProIlelleleAenGlnAenTYrCyAaPargVal 365
Db      1093 ATTGTCAATGACAAAGCTCAAAATCTTAATCAATGATCAATTAATCAATGATCTTCAAG 1152
Qy      366 GluProCyYlleGlnIlePheSerAlaValaGlnValYsaenValYrGluAenIle 385
Db      1153 GACCATGTCGGAACAGAAATCAGCTGTGAAGTGAAGCAATCTTCCATGAACATC 1212
Qy      386 lySglYThraSerAlaThraYsaValaAlaIleYsaPheAapCySerThraenPheProCy 405
Db      1213 CGTGAACAAAGTGCCTGAGGTGGCTGTAAATTCGATTTGACAGAAAGTCTCTCATGC 1272

```

Qy	406	GlUGlYtIlelleMeGClAsnllleAsnleuValGlYglUse-rGlyLysProSeGluLa	425
Db	1273	CAAGGAATATTGTGTGGAAACATCAATCGTTGGGATGAGGAAAGGAACACATCATG	1332
Qy	426	ThrCysLysAsnVal	430
Db	1333	TCATCGACGAAATATT	1347
RESULT 13			
LOCUS	AY078936	1364 bp	mRNA linear
DEFINITION	Arabidopsis thaliana At2g41850/T11A7.5	complete cds.	
ACCESSION	AY078936		
VERSION	AY078936.1	GI:19310404	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Shim, P., Chen, H., Cheuk, R., Kim, C. J., Meyers, M. C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Wu, H. C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.		
TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1364)		
AUTHORS	Shim, P., Chen, H., Cheuk, R., Kim, C. J., Meyers, M. C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Wu, H. C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-FEB-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
FEATURES			
source	Location/Qualifiers		
	1. 1364		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="2"		
	/clone="At2g41850.6 (S500052) "		
	/ecotype="Columbia"		
	50..1351		
	/note="putative polygalacturonase"		
	/codon_start=1		
	/product="At2g41850/T11A7.5"		
	/protein_id="AA084942.1"		
	/db_xref="GI:19310405"		
	/translation="MARCTNLVTFLLMLMLMSWCASRISPNVYDHSYKRKSLDKLRREITIGRSFVARSLEPTVSVSDGAGDKDPTDQAFVAAKKACSSNANLVLPKGNITLKSILDTGPGNSILTVQIFGTLASAOKSDYDIDSKMTIMFDSVNLSTV DGGDVGUNGEGETWQNSCKRNKAKICTYAFATLFFINSKSLIVKRLKVNAAQOIQI SIEKSNVQSVNSVVTAPADSPNTDGKIHITNTQNTRVSESLIGTDDICISLSEGQNY QINDITGCGEHSIGSLGDDNSKAFVSGVYDAGLSTGDNGVAILKYQGGSGTASN ILFQNIQMDNVNPIIIDQDCDDKCKTEKSAVQKVVVYDIDSGTSASENALTFNC SSKNYCOOGVILDRVNIKGGKATCTNANVVDKAVLPQGNST"		
	104		
misc_difference	/note="compared to genome"		
	/replace="c"		
misc_difference	445		
	/note="compared to genome"		
	/replace="t"		

	misc_difference 1124	/note="compared to genome"
	misc_difference 1129	/replace="g"
		/note="compared to genome"
		/replace="g"
ORIGIN		
Alignment Scores:		
Pred. No.:	2.13e-87	Length: 1364
Score:	1149.00	Matches: 226
Percent Similarity:	67.71%	Conservative: 76
Best Local Similarity:	50.67%	Mismatches: 118
Query Match:	48.08%	Indels: 26
DB:	8	Gaps: 6
US-10-691-374-2 (1-457) x AY078936 (1-1364)		
QY	6 AsnSerIleLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAen 25	
Db	65 AACCTTGTAACGGCTTCTCTATTAGGGCTCTTTGATGCTCAGTGTAAAGCTTCA 12	
QY	26 ValIleAspAspAsnLeuPheLysGlnValTyr-----AspAsnIleu 40	
Db	125 AGAAATCAGCCCTTAATGATATGACCATCTTATTAAGTTCAATCCGATAGCTTAATC 18	
QY	41 GluGlnGluPheAlaIleAspPheGlnAlaTyrLeuSerTyrLeuSerIysAsnIleGlu 60	
Db	185 AAGCAAGA-----GAGGACATACGCGAGTTGAAGAGTTGTAAGACCTCTTGGCG 23	
QY	61 SerAsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSer 80	
Db	239 ACTCCAAAC-----ACGTTAGTGTCTGAT 26	
QY	81 PheGlyAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGluGlnAlaTrpAsn 10	
Db	266 TTTCGAGCTTAAGAGATGAGAAACCGATGACGACGCGCTTCGTGAATCGTGGAG 32	
QY	101 GluAlaCysSerSerSerArgThrProValGlnPheValIleProLysAsnLysAsnTyrLeu 12	
Db	326 AAAGCATGCTTCTCAATGAGCGGTGAATCTCTAGTTCCTTAAGCGAATCTTATCTC 38	
QY	121 LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValIlePheGly 14	
Db	386 CTTAAGCTTATTCATTAATTAACGTGTCATGCAATTTCTATTCACCGTTCAGATATCTCGT 44	
QY	141 SerLeuGlnAlaSerSerSerLysIleSerAspTyrLysAspArgArgLeuTrpIleAlaPhe 16	
Db	446 ACGTTATCCGATCTCAAAACAGATCGATTACAAAGATATACGCAAAATGATATGTTT 50	
QY	161 AspSerValGlnAsnLeuValAlaGlyGlyIly-----GlyThrIleAsnGlyAsnGly 17	
Db	506 GATGGCGTTTAAACATCTATCATGTCGATGAGGCGACACATGGGGGTGTGAGCAAAACGCC 56	
QY	179 GlnAlaTrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaPro 19	
Db	566 GAACCGTGGCGCAAAATCTATGCAAAACGGAACAAAGCTAAACCATGACCAAAAGCCCCA 62	
QY	199 ThrAlaLeuThrPheTrpAsnCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsn 21	
Db	626 ACGGCTCTTACTTTTAAACAATCGAAAGAGTGTGATGAGGAAGATCTGAAGGTGAGAAAT 68	
QY	219 AlaGlnGlnIleHisIleLysPheGluSerCysThrAsnValAlaAlaSerAsnLeuMet 23	
Db	686 GCACAGCAGATCCAGATTTGCAATTGAAAGATGCTCAACGTTACGCTCTCAATGTCGTG 74	
QY	239 IleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyr 25	
Db	746 GTAACGCGCGCGGATGTCCTTAACACCGATGATATCTATATCACTAACACCCAAAC 80	
QY	259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly 27	
Db	806 ATTGAGGTCTCGAATCCATCTTGTGAACAGGCGATGTTGCATATCTATTGAAAGTGA 86	

Oy		279	SerGlnaenValIGlnAlaThrAsnLleThrcysGlyProGlyYHlAGLYlleSerTleGly	298
Dd		866	TCAAAATGTTCCAATCATGATATACTGGCGGTCCGGGCACGGATACGATTATGGG	925
Oy		299	SerLeuGIysSerGIyaenSerGIuaLaTyArVaISeraenValIthrVaIamGIuaLaLyS	318
Dd		926	AGCCTTGGAAGATGACAATTTCCTTAAGCGCTTTGTCTCAGGCGGTGACTGTGGATGGCTAAG	985
Oy		319	IleIleGIyAlaGIuaenGIyValINrgIleLySThrTPgInGIyLySerGIyGlmaLa	338
Dd		986	CTTTCGCCGTACAGCAATGAGTAGAATCAAAACTTTCACAGGAGGGGTCAAGAACTGCT	1045
Oy		339	SerAnIleLyPheLeuAsnValGIuuecGIuaenValLySTyrProIleIleLeap	358
Dd		1046	AGCAATATTAATATTCCAACATTCAGATGATATGTTAAAGATCCACATCATANTGCAC	1105
Oy		359	GlnAsnTyCySaAParVAlGIuProCysIleIngInPheSerAlaValIGlnValLyS	378
Dd		1106	CAMAGCGACTCGACAAAGCGCAA---TGCACTACAGAGAATCCGGCGTCCAAAGTGAAG	1162
Oy		379	AenValValTYrGIuaenIleLyegLyThreSerAlatTrLySValAlaIleLeypHeap	398
Dd		1163	AACGGGTGTATCCGGGACATAGTGCACCTGCGCATCGGAAMAAGCATATACGTTTAAAC	1222
Oy		399	CysSerThrAnpPheProCysGIuGLyIleIleImetGIuaenIleAsnLeuValGIyGU	418
Dd		1223	TGCAGCAAGAACTATCCATGCCAGAAATTTGTCTTGACAGAGTACATTAAGGA---	1279
Oy		419	SerGIyLyPProSerGIuaLaThrCYeLySaSnValIHlApheaSnAmalagluHIeVal	438
Dd		1280	-----GGMAAGCAACTTGACCAATGCTATATGTGTGATTAAGAGAGCTGTT	1327
Oy		439	ThrProHIsCyGethSer 444	
Dd		1328	CTGCCTCAGTGCACACTCC 1345	
<b>RESULT 14</b>				
PAPOLYU				
LOCUS	PAPOLYU	1755 bp	mRNA	linear PLN 11-MAY-1995
DEFINITION	P.americana mRNA for polygalacturonase.			
VERSION	X66426			
KEYWORDS	X66426.1 GI:22630			
SOURCE	polygalacturonase.			
ORGANISM	Persea americana (avocado)			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; magnoliids; Laurales; Lauraceae;			
	Persea.			
REFERENCE	1 (bases 1 to 1755)			
AUTHORS	Doplico, B., Lowe, A.L., Wilson, I.D., Merodio, C. and Grierson, D.			
TITLE	Cloning and characterization of avocado fruit mRNAs and their			
JOURNAL	expression during ripening and low-temperature storage			
MEDLINE	Plant Mol. Biol. 21 (3), 437-449 (1993)			
PUBMED	93184201			
REFERENCE	8095163			
AUTHORS	2			
TITLE	Doplico, B.			
JOURNAL	Direct Submission			
	Submitted (27-May-1992) B. Doplico, Univ. of Nottingham, Dept. of			
	Physiology & Envir. Sciences, School of Agriculture, Sutton			
	Bonington, Loughborough, Leicestershire, LE12 5RD, UK			
FEATURES	Location/Qualifiers			
SOURCE	1..1755			
	/organism="Persea americana"			
	/mol_type="mRNA"			
	/cultivar="Haas"			
	/db_xref="taxon:3435"			
	/clone="SBA24"			
	/tissue_type="pericarp"			
	/clone_lib="cDNA"			
	/dev_stage="ripening fruit"			
	49..1437			
CDS				

		/EC_number="3.2.1.15"	
		/codon_start=1	
		/product="polygalacturonase"	
		/protein_id="CAA47055.1"	
		/db_xref="GI:22631"	
		/db_xref="GOA:002096"	
		/db_xref="UniProt/Swiss-Prot:Q02096"	
		/translation="MALTRLLPISILRCPGSSHTTLOKDDLLCVNGDPPDDOBA /translatability=1 PEKMKADCSGSSVLIIVENKNYLIKQTTPSPGPKSDLVKRTGTEASDDSDWQDK NRKWIIEFEDISNLTEKCGTASNGEETWMSCKRKSLPKSAPFALFRSCKNL VSDLSIKDSQKMLHSFDCQDVIASNLMTAPESHPTGHIITGFORLHVMSVIG GDCCISIESSGSKMVIATITCGPHGISIGSIGDSSEKSHVGLVDGNLPFTTGK RIKTKWSSGSSKNVATKRONIVMHTNPNIIDDOYCDSDPDEPESAKVSNVAIWM IRGTSASEVAIFDCKSSPCGGYIVGININLVANGKETTMSCSNIVGLREGLST LFMRKRVHECSY"	
	ORIGIN		
	Alignment Scores:		
	Pred. No.:	3,59e-86	Length: 1755
	Score:	1136.00	Matches: 223
	Percent Similarity:	71.60%	Conservative: 67
	Best Local Similarity:	55.06%	Mismatches: 107
	Query Match:	47.53%	Indels: 8
	DB:	8	Gaps: 5
	US-10-691-374-2 (1-457) x PABOLYGU (1-1755)		
QY	28	ASPAEPANLEUPHELYGGINVALTYRAEPANILELUGLUGINGLUPHEALAHIAEP	47
DB	166	GACCAAGGGCTACCCACCATTTTGGTCCCATTTTGAAT---GAGTTCTTACATC	222
QY	48	PHGINALATYLEUSERTYLEUSERLYBAENILEGUSERASNAENILEAPLYS	67
DB	223	ATGGGATTGAGCCCTCCATCTTGAAT-----TTGACGCGGTTTAAATCAGGCGTGCT	276
QY	68	VALAPLYBAENGLYILEYVALILEANVALLEUSERPHGELYALASGLYAPGLY	87
DB	277	CCGGAAGATTCCAGATACAGATTTAAAGTGTGATGATTTGGACCCAGAGGTATGGG	336
QY	88	LYSTHTYRAEPANILEALAPHEGLUGINALTPANGLUALACYSERSETATGTHR	107
DB	337	---ACAGATGACCAAGGCATTGAGAGGCTTGGAGAGCGCTTGTTCATAGATCT	353
QY	108	PROVALGINPHEVALIAPROLYSASNLYBAENITYLEULYUGLINLETTPHESER	127
DB	394	GTCCTC-----ATTGTGCTGTAAACAAAGACATATCTTAAACAAATTAACCTTTCA	447
QY	128	GLYPROCYBARGSERSETLESERVALYVILIEPHEGLYSERLEUGLUALASERLYS	147
DB	448	GGTCTTGTAAATCTGATCTTGTGTGAAGATTCGCGGAACATGMACTTCATCTGAT	507
QY	148	ILESEAPIRY-----LYSAPRAGTATGULEUTPILIALAPHEAPSERVALIGIAN	165
DB	508	CAATVALAGITGGGTGGACATATGAAAAACATGATGAAATTTGAAGACATACAGAAC	567
QY	166	LEUVALIAGLYGLYGLYTHRIIEANGLYAENGLYGLNVALITPTTPROSERSER	185
DB	568	CTCAACACTAGAAAGGGGTGGACATCCCATGGAATGAGAGACATGGTGGATACCTCC	627
QY	186	CYSILYVILEANLYSEITLEUPROCYBARGAEPALAPROTHALALEUTHPHETTPHAN	205
DB	628	TGCANAAAGAAATCCCTTCTTGCANAAACCGCACCTACGGCTTAACTTCAAGTCC	687
QY	206	CYSILYBAENILEYVALIASNAENILEYSSERYBAENALGLINGLINLEHISILEY	225
DB	688	TGCAGAAACCTGATTTGTAGCATCTCACTATCAAGAGACGCCAAAGATGCACTTTCT	747
QY	226	PHGLUSERCYETHASNAVALVALASERANLEUMETILEASNALASERALALYSER	245
DB	748	TTTGAITAGTGCAGAGATCTCACTCTTAATCTCACTGTAAGTCACAGAGCATAGC	807
QY	246	PROANHTHAPGLYVALIHIVALSERANHTHNGINTYRILEGLINLESERAPHTHRIE	265

[illegible]

```

gene
CDS
1..1765
/gene="pegaz"
73..1317
/gene="pegaz"
/codon_start=1
/product="endopolygalacturonase"
/protein_id="CAC05658.1"
/db_xref="GI:9967520"
/db_xref="GOA:O9FDY7"
/db_xref="UniProt/T/EMBL:O9FDY7"
/translation="MAHCCTVLTGLLWMSLFMLSCWCAARNIRKRDHSTYGFENSGI
IKRRDDITRLKSVARASLRPLTYSVSPFGAKGCTDDTDAFNAAMKACSSGAVN
LVLVEGKTIFLKSLRLNGPCKSVLTVOILGLTSAQQRSDYEDLSKMTIFGVNSLTIV
DGGATGVNNGAGETWMWNSCKRNEAKCTAPALPFPYNSKILRNVLNRKYDAQQQLC
ISIEKSNVQSNVNEVTPADSPNPTDGIHITNTNIGVNSNIIGTGDDCISIEEGSONI
VOINDLTCVGPBGHSIGSLGDNBSKAVSGVYDAGALGTGDNVRLIKTNQSGGTSAS
NIIPQNTIOMVENVDPRIIDQYCDKSKCTEKSASVQIKYNNVRIISGTSASDIATFNR
CSKNYPCGATLDEKVINCGKASCSNNANMDKQVLEOCKST"

```

<b>Alignment Scores:</b>					
Pred. No.:	1,16e-85	Length:	1765		
Score:	1130.00	Matches:	221		
Percent Similarity:	67.34%	Conservative:	78		
Best Local Similarity:	49.77%	Mismatches:	117		
Query Match:	47.28%	Indels:	28		
DB:	8	Gaps:	7		

  

US-10-691-374-2 (1-457) x BNA250919 (1-1765)

QY	8	IleLeuLeuIleIleIlePheAlaSerSeriIleSerThCyArgSerAsnValIle	27
Db	103	TTACTAGTACGCTCTTTTATTANGCTCTCATGGTGGAAGCTTCAGAAACAACATTT--	15
QY	28	AspAspAsnLeuPheUySGlnValTyraSpAsnIleLeuGlInglInPheAlaHisasp	47
Db	160	-----AATAGATATGATGATCTCT	17
QY	48	PheGIlnAlatyrLeuSer---TyrLeuSerLySaenIIegLusEraShsbnsbnIIeaSp	66
Db	178	TATGAAAGCTTCAAATCCATAAGCTTAATCAAGCGAAGAGATATACGTCTGTGAAA	23
QY	67	LysValaAspyrSaenglyIleUySVaI-----IleaSnValLeuSer-PheGIyALA	83
Db	238	AGTCTTCAGAAGACCCTCTTGCGGCTCCAACCCGTTTAGTGTTCGATTTTGAACT	29
QY	84	LyeGIYAapgyLySrThTyrrAspAsnIlealapheluglInalatriPaSngluAlaCys	103
Db	298	AAGAGAGATGAAAAACCGACGACACGAGCGTTTCGTGAATGATGAGAGAAAGCTTCT	35
QY	104	SerSerArgrThrProValGlnPheValAlaPolySaSnLySaenTyrrLeuUlygLn	123
Db	358	TCCTTCAGTGCAGCTTMAATCTCTTAGTTCCTGAAGGAAAAACTATTTTCTCAAGTCT	417
QY	124	IleThrPheSerGIyProCYsarGerSerIleSerValLySIlePheGIySerLeuglu	143
Db	418	CTTCGATTAAATGGCCCATGCAAACTOAGTCTTACCGTTCAGATTTTAGGAACCTTATCG	477
QY	144	AlaSerSerLySIleSerApTyrrLySaSPArGrLeuTriPIlealaphaeASpSeVal	163
Db	478	GCTTCTCAACACGATCGGATTATGAAGTCTCAGCAAAATGGATTAGCTTGAAGGGGTT	537
QY	164	GlnAsnLeuValAlglyIglyI-----GlyThrIleanglySaengly--GlnVal	180
Db	538	AACAGTCTAACGGTTGATGGCGAGCACCGGAACCTGCAACGMAACGGGGCGGAACG	597
QY	181	TrIPtrPrSerSerCyelysaIleasnlySSerLeuProCYsarGaSPAlaproThAla	200
Db	598	TGATGGAAAACTCATGCCAACGGAACGAGACTTAAGAAATGCACAAAGCCCTCAGGCT	657
QY	201	LeuThrPheTriPasnCylLySaSnLeuLyValaSnAsnLeuLySSerLySaSnAlaGln	220
Db	658	CTTACTCTTCAACACCGGAAGATTATTAATGATATATCTGAAGCGGAGGAAGATTTCCA	717

```
QY 221 GlnIleHisIleIyepheGluSerCysThrAsnValValAlaSerAsnLeuMetIleAsn 240
    |||||
DB 718 CAGATTGAGATTTCGATTGAGAAATGCTCAACGTTCAAGTCTCAAAATGCTGAGGTCACCT 777
QY 241 AlaSerAlaIySerProAsnThrAspGlyValHisValSerAsnThrGlnTrpIleGln 260
    |||||
DB 778 GCACCTGCTGATGATGCTTAACACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
QY 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280
    :::::
DB 838 GTCTCAATTCCATTCATTCGAAACAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 897
QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300
    |||||
DB 898 AACGTTCAAAATCAACGATTAATGCTGCGGCCAGGCCAGGATGATGATGATGATGATGATGATGATGAT 957
QY 301 GlySerGlyAsnSerGlyAlaTrpValSerAsnValThrValAsnGlnAlaIySileIle 320
    |||||
DB 958 GGAAGATGACAAATTCAAAGGCAATTTGATCAAGCGTGACTGTGATGATGATGATGATGATGATGATGAT 1017
QY 321 GlyAlaGluAsnGlyValIyrgIleIythrTrpGlnGlyGlySerGlyGlnAlaSerAsn 340
    |||||
DB 1018 GGTACAGACAAATGAGATTGATGATCAAAACTTAACACAGGAGGTTCTGGAATGCTGATGACAAAC 1077
QY 341 IleIyepheLeuAsnValGluMetGlnAspValIySerTrpIleIleIleAspGlnAsn 360
    |||||
DB 1078 ATTATATTTTCAGAAACATTCAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1137
QY 361 TyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValIyAsnVal 380
    |||||
DB 1138 TATTGCGACAAG---AGCAAAATGCACTGAAACAAAATCAGCGGTCCAAATTAAGAAAGTGTG 1194
QY 381 ValIyrgIluAsnIleIyGlyThrSerAlaThrIyValAlaIleIySerPheAspCysSer 400
    |||||
DB 1195 GTGTACCGGACATTAAGCGGCACAGCGCATGCGACATGACATTAATCTTCAACTGCGACGC 1254
QY 401 ThrAsnPheProCysGluGluIyIleIleMetGluAsnIleAsnLeuValGlyGluSerGly 420
    |||||
DB 1255 AADAACATATCCATGCAAGCAATGTACTGACAAAGTGACATTAAGGGA----- 1305
QY 421 LysProSerGluAlaThrCysIyAsnValHisPheAsnAsnAlaGluHisValThrPro 440
    :::::
DB 1306 -----GGGAAAGCCGATGTCAGCAATGCTAATGTGATGATGATTAAGACCGGTGTTTACTT 1359
QY 441 HisCysThrSer 444
    |||||
DB 1360 CAGTGCAAATCC 1371
```

Search completed: March 5, 2005, 00:17:23  
Job time : 5020 secs

**This Page Blank (uspto)**

**This Page Blank (uspto)**